UNIVERSITY OF LONDON THESIS

Degree PhD Year 2005 Name of Author MARSHALL, C.M.

The morpho-phonological interface in Specific Language Impairment

Chloe Ruth Marshall

Thesis submitted in partial fulfilment
of the requirements for the degree of Doctor of Philosophy

University College London
October 2004

For my parents, Rosemary and William, who have taught me to appreciate life's most precious treasures – friendship, travel and learning.

Acknowledgements

In October 2000 I gave up teaching in order to study linguistics. On a whim (their website was welcoming, easy to navigate and, in those days, a little old-fashioned), I chose the linguistics department at UCL and spent twelve very happy months there studying for a Masters degree. The inspiring teaching of Neil Smith, Hans van der Koot, Ad Neeleman, Deirdre Wilson, Robyn Carston, Geoff Lindsey, Moira Yip and John Harris meant that I very quickly decided stay on to do a PhD. It was inevitable that I should choose to continue with phonology: the theoretical framework which John has shaped, and which he communicates with such enthusiasm, is elegant and satisfying.

I have been very privileged in being able to carry out my doctoral research at the Centre for Developmental Language Disorders and Cognitive Neuroscience. Huge thanks are due to Heather van der Lely, Angela Pozzuto, Karen Froud, Theo Marinis, Stavroula Stavrakaki, Elisabeth Fonteneau, Hilary Gardner, Rosalind Hill, Susan Ebbels, Marisa Silveira and Nichola Gallon - their intellectual stimulation and support are much valued. In the Department of Human Communication Science where we are based, Shula Chiat, Ruth Campbell, Jane Maxim, Sarah Alleemudder, Chris Donlan, Chris Fryer and Mike Coleman have provided support of many kinds, and I thank them gratefully. The members of the Phonology Reading Group have provided a very positive critical forum for trying out new ideas and practising conference talks, and I thank in particular John Harris, Moira Yip, Mary Pearce, Gloria Malambe, Nina Topintzi, Irini Sanoudaki, Sam Hellmuth, Alex Bellem and Dorota Glowacka.

Within the wider research community, many people have been generous enough to spend time with me discussing my ideas and theirs. As many of them have been my intellectual heroes for the past few years, it has been a great privilege to meet them and to discover just how warm and genuine they are in person: Jill Beckman, Harald Clahsen, Martha Crago, Katherine Demuth, Susan Gathercole, Cecilia Kirk, Laurence Leonard, Philip Lieberman, Gary Marcus, William Marslen-Wilson, Jim Morgan, Marina Nespor, Steven Pinker, Franck Ramus, Mabel Rice, Yvan Rose, David Sobel, Melanie Soderstrom, Joseph Stemberger, Michael Thomas and Michael Ullman. I had great fun in September 2004 with Katherine Demuth and her team at the Child Language Lab, Department of Linguistics and Cognitive Sciences, Brown University, and thank everyone there for stimulating discussions.

Also generous with their time have been the many children and young adults who have been willing to submit themselves to endless testing during my two years of data collection, often for no reward other than the excitement of missing school lessons (although some did get the occasional sticker). I am grateful to the children and their parents and teachers for welcoming me on my visits.

Much as I have loved working on this thesis the past three years, I have not let it take up all my time. Time spent away from studying has been a source of great pleasure, and for that I thank my family and friends. In particular, my mum, dad and granny have concocted copious cultural, culinary and conversational distractions, whisking me off to museums and concerts, lending me novels, plying me with delicious food and wine, and reminding me that psycholinguistics is not the only topic worthy of discussion. With my friends, time spent holidaying, playing with babies, making merry at dinner parties, watching movies, enjoying walks in the country, and laughing till our sides split has been incalculably precious. I hesitate to name them in case I inadvertently miss someone out, but at the last count they are Claire, Geoff, Lisa, Robin, Charly, Iain, Juliet, Karina, Simon, Kavita, Ted, Rachel, Tom, Helen, Matt, Nikki, Sue, Duncan, Mark, Pat, David, Jackie, Barry, Jack, Pat, Jon, Michelle, Vikki, Beate, Tim and Alison.

My greatest debt is to Heather. In my second week at UCL I attended a seminar that she gave at the Department of Phonetics and Linguistics, and immediately knew that I wanted her to be my thesis supervisor. I'm still not sure why she agreed to take me on, because I can't imagine I made a very good first impression - in fact, I know I didn't! I hadn't read any of her papers and knew very little - actually, nothing at all - about SLI, but managed to convince her, somehow, that I was willing to learn. My first piece of written work for her was atrocious, and yet she didn't give up on me. Heather has been the very best of teachers, in that however high the expectations I have had of myself, she has encouraged me to aim even higher. Her faith in me, her encouragement and patience, her enthusiasm for my ideas and her willingness to give up so much of her time for me are, I hope, repaid in some small way by the work that follows.

Abstract

This thesis investigates the nature of the interface between two components of language – morphology and phonology – in children with Grammatical-Specific Language Impairment (G-SLi), compared to those with typically-developing language. I focus principally on the impact of phonological complexity on past tense inflection, but I also investigate other areas of morphology. More specifically, I show that for G-SLI children:-

- There exists a phonological impairment that is independent of morphology. This impairment is characterised by the simplification of complex syllable structure, and by syllabic and segmental errors when the word starts with an initial unstressed syllable.
- There exists an impairment in past tense morphology, characterised by suffix omission, that is independent of phonology.
- Phonological factors affect past tense morphology. Specifically, suffix omission rates are higher when inflection (i) creates clusters at the word-end or (ii) requires the syllabic allomorph /ɪd/.
- Phonological factors also affect plural and present progressive formation.
- Unlike past tense morphology, derivational morphology is not subject to suffix omission. However, non-target derivational forms result when stimuli are morphologically or phonologically complex.

I argue that grammar has a modular structure, and I propose that deficits in one or more of the following modules – syntax, morphology and phonology – can impact on past tense inflection. This model, termed the 'Computational Grammatical Complexity' (CGC) hypothesis, can account for why tense is an area of exceptional difficulty for children with SLI.

This investigation is underpinned by a rigorous theoretical framework. Not only does using a cognitive scientific and linguistic framework further our understanding of the nature of the deficit in SLI, but SLI provides a valuable testing ground for theories of language acquisition and the representation of language in the brain.

PART 1. THE BACKGROUND

PART 4. INVESTIGATING THE IMPACT OF A PHONOLOGICAL DEFICIT ON OTHER AREAS OF MORPHOLOGY

PART 1.

THE BACKGROUND

Chapter 1. Introduction

1.0. Chapter outline

This chapter sets out the issues under investigation, and the theoretical framework that I adopt in this thesis. I begin by discussing the nature of SLI and the two opposing viewpoints that have been proposed to account for this disorder (Section 1.1). I then set out the linguistic- and cognitive-theoretic frameworks within which I analyse my findings, and use these frameworks to discuss previous studies of the acquisition of phonology and morphology in both typically developing and SLI children (Section 1.2). Next I consider the interaction between morphology and phonology in typical development and SLI (Section 1.3). I finish (Section 1.4) by setting out the structure of the thesis.

1.1. The nature of SLI

1.1.1. What is SLI?

Children with Specific Language Impairment (SLI) have significantly impaired language acquisition despite the absence of any obvious language-independent cause, such as hearing loss, low non-verbal IQ, motor difficulties or neurological damage (Leonard, 1998). Furthermore, the impairment is noticeable at the outset of language development: it does not emerge in later childhood as the result of some sort of trauma or illness, (Bishop, 1997; van der Lely, Rosen & McClelland, 1998). It is estimated to affect approximately 7% of the English-speaking pre-primary school population, and a significant proportion of these children experience severe and persistent difficulties that impact significantly on school and career attainment (Leonard 1998).

Within the SLI population as a whole, deficits have been diagnosed with the 'core' grammatical areas of syntax, morphology and phonology, and, to a lesser extent, the lexicon. Most researchers would agree that syntactically simple sentences, inflectional errors, poor phonological abilities and delayed lexical acquisition are characteristic of SLI (Bishop, 1997; Leonard, 1998). The picture is complex, though, because the range of impairments and their level of severity, stage of resolution and degree of compensation all vary greatly between individuals. The picture is further complicated by findings that children with other developmental disorders, such as Down's Syndrome, Fragile X syndrome and Autism, suffer delays and deficits in their acquisition of language (see Tager-Flusberg, 1999, for a review). The borders between SLI and these other disorders are difficult to define (e.g. Bishop & Norbury, 2002).

SLI has become recognised as a valuable testing ground for teasing apart the relative contributions of domain-general and domain-specific cognitive mechanisms, and for testing modularity within the language system (Levy & Kave, 1999; Pinker, 1991; van der Lely, 1997a,b). However, because SLI is highly heterogeneous, a single explanation is unlikely to be able to account for the broad range of impairments, and this heterogeneity makes it difficult to test linguistic and cognitive models of the disorder. One way out of this impasse is to identify subgroups of SLI children whose members share a common profile of linguistic strengths and weaknesses. In the next section I outline the attempts that have been made to identify different subgroups within the SLI population.

1.1.2. Subgroups of the SLI population

Many researchers have tried to divide the SLI population into subgroups (e.g. Aram & Nation, 1975; Rapin & Allen, 1983), but categorisation has most often been based on clinical rather than linguistic criteria. For example, in the US the 'Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV; American Psychiatric Association, 1994) uses the subtypes 'expressive' and 'receptive and expressive', which give little clue as to the underlying linguistic deficits.

One approach to looking for subgroups is to seek out cases of language impairments that run in families. A rare form of familial language impairment (FLI) is exhibited by the KE family (Gopnik & Crago, 1991; see papers in Matthews, 1994). Affected members of this multi-generational English-speaking family have deficits in many aspects of expressive language, and suffer additionally from verbal dyspraxia, which affects articulation (Gopnik & Crago, 1991). Inflectional morphology, including plural formation (Goad, 1998) and past tense production (Ullman & Gopnik, 1999), is impaired. Derivational morphology is also impaired (Dalalakis, 1994; Gopnik & Crago, 1991). While some affected members of the family have low non-verbal abilities, others score in the average range: non-verbal deficits are therefore not characteristic of the disorder. A gene (FOXP2) implicated in language impairment in this family has been located on chromosome 7 and sequenced (Lai, Fisher, Hurst, Vargha-Khadem & Monaco, 2001). Affected members have a single nucleotide substitution in this gene, which encodes a transcription factor. However, it should be noted that a defect with the FOXP2 gene has only been found in one other individual who suffers from a similar speech and language impairment (Lai *et al.*, 2001), but not in around one hundred other children with SLI and members of their familes (The SLI Consortium, 2002). The FOXP2 mutation is therefore unlikely to be implicated in the majority of SLI cases (Marcus & Fisher, 2003).

A second approach to identifying subgroups within the SLI population is to look for groups of children with relatively uniform linguistic characteristics. Van der Lely and her colleagues have identified a group of children, termed the Grammatical (G)-SLI subgroup, whose difficulties with language appear to be confined to the core aspects of grammar – syntax, morphology and perhaps phonology (van der Lely, 1996a, 1997a,b, 1998; van der Lely *et al.*, 1998; van der Lely & Stollwerck, 1997). G-SLI is consistent with an autosomal dominant inheritance (van der Lely & Stollwerck, 1996). Crucial to its characterisation are the persistence of the deficit (children are aged 9 years and over) and the particular pattern of grammatical impairment.

Within syntax, particular difficulties are evinced when non-local syntactic dependencies such as the use of subordinate clauses (van der Lely *et al.*, 1998), wh-question formation (van der Lely & Battell, 2003), the assignment of thematic roles (van der Lely, 1996a) and the assignment of pronominal and anaphoric reference (van der Lely & Stollwerck, 1997) are required. In terms of morphology, these children omit past tense inflection at high rates (van der Lely & Ullman, 2001) and produce regular plurals inside compounds (van der Lely & Christian, 2000). Although their articulation is intelligible, initial work has revealed subtle deficits in prosody affecting consonant clusters and unfooted syllables (Gallon, 2002; Marshall, Harris & van der Lely, 2003; Peiris, 2000). However, they do not evince non-verbal cognitive impairments, nor consistent auditory impairments (van der Lely *et al.*, 1998; van der Lely, Rosen & Adlard, in press).

In some children, word-finding and vocabulary difficulties appear to be the primary linguistic deficits. A group of children with word-finding difficulties, despite normal non-verbal intelligence and no articulation, hearing or neurological difficulties, has been identified by Dockrell, Messer and George (2001). These children display a wide-ranging profile with respect to other language skills, but show poor performance in tests of accuracy and speed of naming. However, the linguistic abilities of these children were not adequately characterised in this study, and so it is not clear whether these children also have grammatical difficulties. Froud and van der Lely (in prep.) have identified children with impaired vocabulary well below the norm for their age but with relatively good grammatical abilities (for instance, they make few errors of verb inflection). These children score within normal limits on a test of sentence comprehension, but score poorly on measures of expressive and receptive vocabulary. Froud and van der Lely hypothesise that these children have a deficit which primarily affects lexical representations, and they term this group of children the Lexical (L-) SLI subgroup.

Rapin and Allen (1983) described a subgroup of SLI children who have fluent and grammatically well-formed language but a striking inability to comprehend and engage in communicative discourse, and who frequently give inappropriate responses to questions. They termed this the Semantic-Pragmatic subgroup. In a later study, Bishop, Chan, Adams, Hartley and Weir (2000) used discourse analysis on conversational data gathered in a semi-structured situation to identify a subset of language-impaired children with communicative impairments ranging beyond basic deficits in grammar. These children have disproportionate difficulty, in comparison to other SLI subjects, in responding to and expressing communicative intents, and are termed as having Pragmatic Language Impairment (PLI). However, in more recent work Norbury (2003) has claimed that there are no straightforward distinctions between SLI, PLI and High Functioning Autism, and that such labels have no validity.

It should be noted that the existence of subgroups, and particularly the G-SLI subgroup, is a controversial issue (e.g. Bishop, Bright, James, Bishop & van der Lely, 2000; Tomblin & Pandich, 1999). For example, Tomblin and Pandich claim that children with profiles such as AZ's (the child with G-SLI reported in van der Lely, 1997b, and van der Lely *et al.*, 1998) are just part of the normal variation that will be found when comparing grammar and vocabulary scores. They consider it methodologically problematic to present individual cases as evidence for a separation between language subsystems. On the other hand, it is doubtful what studies that do *not* acknowledge the heterogeneity of SLI can actually tell us about the disorder. For example, studies that include children or adults with differing linguistic profiles will produce averaged data that could be camouflaging different patterns of development. Relying on these data will then lead to explanations and predictions that are not generalisable to any particular group of SLI children. This is disadvantageous in both clinically-oriented research and in research which aims to characterise the deficit in linguistic- and cognitive-theoretic terms.

1.1.3. Perspectives on the causes of SLI

There are two main perspectives as to what causes SLI. The first is the domain-general perspective, which holds that an input-processing deficit interferes not only with various aspects of language acquisition, but also with the acquisition of non-linguistic cognitive skills. The second is the domain-specific perspective, which claims that the deficit is specific to grammar and independent of non-linguistic skills. These perspectives are in turn related to the larger debate of how the brain is organised, and how specialised cognitive systems such as language develop. Some researchers claim that general-purpose mechanisms become specialised through experience during development, and therefore contend that pure impairments of a specialised system such as language cannot exist (e.g. Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett, 1996; Karmiloff-Smith, 1998).

Others argue that genetically determined specialised mechanisms underlie different cognitive abilities, including language, and therefore predict that pure primary impairments of specialised systems will exist (Fodor, 1983; Pinker, 1991, 1999).

Because one of the main focuses of this thesis is inflectional morphology, and because inflectional morphology, in particular morphology for tense, is one of the most characteristic deficits in SLI, I discuss the domain-general and domain-specific perspectives with reference to inflectional morphology.

1.1.3.1. The domain-general perspective

The domain-general perspective argues that an input-processing deficit underlies processing abilities that are general to cognition as a whole rather than specific to language (e.g. Kail, 1994). However, it should be noted that although low non-verbal IQ scores are reported in some SLI children (e.g. Bishop *et al.*, 2000) and in some members of the KE family (Vargha-Khadem, Watkins, Fletcher & Passingham, 1995), not all SLI children have low IQ. Nor is there a direct correlation between language abilities and IQ score (Rice, Tomblin, Hoffman, Richman & Marquis, 2004). There is therefore no support for the view that a general cognitive impairment necessarily underlies SLI. Of course, this does not mean that non-verbal IQ *never* affects language skills, and nor does it rule out the possibility that poor verbal abilities impact in some way on non-verbal IQ.

A more specific claim is that a deficit in processing rapidly-presented sequential information impacts on the processing of acoustic stimuli (Tallal & Piercy, 1973a,b), resulting in a phonological processing deficit that hinders language development (Joanisse & Seidenberg, 1998, 1999, 2003; Leonard, 1989, 1998; McClelland & Patterson, 2002).

If an auditory processing deficit were responsible for SLI, we would expect to see evidence that (a) all children with SLI have an auditory deficit, (b) the extent of the deficit correlates with the impairment in language abilities, and (c) no child with normal language has such a deficit (see Ramus, 2003, for similar arguments regarding developmental dyslexia). Van der Lely *et al.* (in press) address all these issues in their study of auditory processing in the G-SLI subgroup. Children were tested on the discrimination of speech, non-speech and tone stimuli, presented at varying rates. 31% of the G-SLI children showed normal auditory processing for speech, 69% for non-speech and 46% for tones. No relationship was found between auditory and phonological/grammatical abilities in these children. Furthermore, several control children failed these tasks, despite their normal language skills. The evidence is therefore not consistent with G-SLI being caused by a deficit in processing rapid acoustic information, although it does not rule out the possibility that this is the case for some children with other forms of SLI. Nor does it rule out the possibility that these children suffered from a perceptual deficit earlier in development, which has since resolved: only longitudinal studies will be able to probe that particular issue.

Within the input-processing deficit perspective, one hypothesis, the Surface Hypothesis (Leonard, 1989, 1998), has been particularly influential. It proposes that the auditory properties of English grammatical morphemes cause difficulty for SLI children. Underlying the Surface Hypothesis is the assumption of a domain-general processing capacity limitation. The argument is that for a cognitive system already under strain due to a deficit in processing ability, the lack of perceptual saliency of grammatical morphemes makes them more difficult to process than is the case for a non-impaired child. This in turn leads to difficulty in setting up morphological paradigms (in the sense of Pinker, 1984), affecting both comprehension and production. The prediction is that incomplete processing results in inflections and function words being omitted. Grammatical constructions such as the passive, which rely on identifying morphemes of brief duration (e.g. *The girl was kiss**ed** **by** the boy*) are affected, as is lexical learning, where surrounding closed-class morphemes aid in the identification of grammatical category (e.g. *the* suggests a noun and *-ed* a verb). Importantly, the Surface Hypothesis assumes no fundamental deficit in the underlying grammar independent of slow intake and processing of incoming language data.

Of critical importance to Leonard's hypothesis is the interaction between morphology and phonology. It is not that the perception of, say, word-final */d/ per se* that is difficult for SLI children. Rather, it is because these children's limited processing ability is taxed when this sound plays a morphological role that there are repercussions for a word with a past tense inflection, such as *stayed*. In this case they have to relate *stayed* to *stay*, hypothesise that */d/* is a morpheme in its own right and place it in the correct morphological paradigm. A monomorphemic word such as *staid*, in contrast, does not pose these difficulties. The Surface Hypothesis leads us to predict that SLI children will have more difficulties in perceiving and producing segments with low perceptual saliency in the context of a word plus a grammatical morpheme than in a monomorphemic context. Although the concept is central to the Surface Hypothesis, 'perceptual saliency' has not been adequately defined in phonetic terms. For Leonard (1989), relative duration is probably the most important relevant property. In English at least, grammatical morphemes are most often single consonants (e.g. past tense */t/* and */d/*, plural or third person singular */s/* and */z/*).

The Surface Hypothesis predicts that two grammatical morphemes with identical phonological patterns (e.g. English noun suffix *-s* and third person agreement *-s*) will behave identically. However, it also predicts that for two inflections with identical phonemes that differ only in relative duration, the inflection that is longer in duration will be acquired more easily. Hsieh, Leonard and Swanson (1999) have shown this for the plural suffix *-s* and the third person singular suffix *-s*. The plural inflection is on average longer in duration than the verb inflection because it is more frequently found in sentence-final position, where it undergoes lengthening. Hsieh *et al.* suggest that these perceptual differences may contribute to the fact that children acquire the noun inflection before the verb inflection. However, this is not the only interpretation: as the authors acknowledge, these inflections differ both semantically and syntactically.

Because the Surface Hypothesis was formulated on the basis of the auditory properties of English morphemes, Leonard has long recognised the need to test it cross-linguistically. In languages where the phonetic realisation of grammatical morphemes is different from English, one would predict a different pattern of impairment for morphemes with identical grammatical function. This is a very different prediction to that made by a model where SLI is caused by an underlying language-specific deficit. The language-specific deficit predicts that grammatical morphemes which share morphosyntactic properties will show a similar pattern of impairment cross languages regardless of any differences in phonetic realisation.

After English, the language which has been most studied in SLI is Italian. Italian, in contrast to English, has a rich inflectional system: all nouns and adjectives are inflected for gender and number, and all verbs have to be inflected for tense, person and number. All inflections are word-final and syllabic, containing minimally one vowel. Most Italian words are multisyllabic, with primary stress falling on the penultimate syllable. Consequently, grammatical inflections are usually word-final weak syllables that are immediately preceded by a strong syllable, such as *cammino* ('I walk') and *ascolta* ('he listens'). The pattern of grammatical morpheme deficits seen in English-speaking children with SLI does not appear to apply to Italian SLI children. The generalised finding (summarised in Leonard, 1998) is that Italian speakers with SLI have very few difficulties with verbal inflection and no difficulties with noun and adjectival inflection. This result is explained as being due not only to the greater perceptual saliency of Italian inflection, but also due to the child focusing his limited processing resources on inflection, as this is much more critical for sentence interpretation in Italian than it is in English. English-speaking SLI children, however, concentrate more on word order, as word-order is more important for sentence interpretation in English than it is in Italian. English SLI children therefore have greater deficiencies in inflection (the 'Sparse Morphology Hypothesis').

Italian SLI children do, however, have difficulty with articles and clitics: this is explained as being the result of these words occurring in unstressed positions, resulting in lower perceptual saliency. Leonard, Sabbadini, Volterra and Leonard (1987) showed that Italian-speaking children with SLI omit preverbal direct object clitics (e.g. *lo*, 'him') much more frequently than English-speaking children with SLI omit the direct object pronoun *him*. One of the differences between these two morphemes is their prosodic position – in Italian, these clitics consist of a weak syllable that usually precedes a finite verb and follows a weak syllable. In English, the pronoun is a weak syllable that is more likely to immediately follow a strong syllable, and is able to occur in clause-final position, where it is subject to lengthening.

Not all cross-linguistic studies are supportive of the Surface Hypothesis. Stavrakaki & van der Lely (submitted) tested the comprehension of pronouns in Greek. They found that SLI children have no difficulty comprehending pronouns that are interpreted on semantic grounds, be they perceptually salient or non-salient. Nor do they have difficulty comprehending non-salient pronouns that are interpreted through local syntactic relations. They do, however, have difficulty with non-salient pronouns that require interpretation through non-local syntactic dependencies. This evidence suggests that it is not the phonological properties of pronouns that determine successful comprehension, but rather their syntactic properties.

1.1.3.2. The linguistic deficit perspective

In direct opposition to input-processing deficit accounts of SLI, such as the Surface Hypothesis, are accounts that propose a deficit specific to the grammar. These accounts are rooted in Generative linguistic theory (Chomsky, 1965). Chomsky proposes that children are only able to acquire language because they possess innate knowledge, a 'Universal Grammar', which guides their learning (for a review, see Pinker, 1994). Researchers studying SLI within this framework have proposed not only that the deficit is located within grammar, but that such a deficit offers direct evidence for the innateness of grammar (Gopnik, 1997; van der Lely, 1997b; van der Lely *et al.*, 1998).

Most accounts of a language-specific deficit for SLI focus their attention solely on explaining difficulties with inflectional morphology. They claim that syntactic features which mark inflection are either missing altogether – as in the Agreement-Deficit Hypothesis (Clahsen, 1989; Clahsen, Bartke & Goelner, 1997) and the Missing Features Hypothesis (Gopnik, 1990; Gopnik & Crago, 1991) - or develop much later than normal - the Extended

Optional Infinitive Hypothesis (EOI; Rice, Wexler & Cleave, 1995). The Agreement-Deficit Hypothesis, as its name suggests, claims that individuals with SLI have difficulty in establishing the structural relationship of agreement between subject and verb because non-interpretable agreement features of verbs are absent or underspecified. According to the Missing Features Hypothesis, SLI grammar lacks sublexical features (e.g. tense, person and number) that mark morphological inflectional information. Consequently the grammar does not contain the morphological rules that introduce these features, and, because of this, there is no internal structure to words. The EOI Hypothesis claims that SLI children experience a prolonged period in which both inflected and uninflected stems are acceptable forms in their grammar, 'extended' because an optional infinitive stage is also characteristic of younger, typically developing, children.

All these hypotheses can account for some of the findings of SLI research, but none is wide-ranging enough to cover the full range of morphosyntactic difficulties observed. For example, the EOI and Agreement-Deficit accounts cannot explain errors other than those with verbs (although a more recent version of the EOI, the AGR-TNS Omission Model (ATOM) claims to explain case errors on pronouns too (Wexler, Shutze & Rice, 1998)). The Agreement-Deficit and Missing Features Hypotheses cannot account for the fact that when tense or agreement features *are* used on a verb, they are used correctly. Although these accounts differ in the fine detail, they all agree that the deficit lies in syntactic features which are either intrinsic to lexical items or added prior to the item entering the enumeration (in the sense of Chomsky, 1995). Certainly there is no suggestion that the deficit is in syntactic operations themselves. This means that none of the accounts can either predict or provide a satisfactory explanation for the range of syntactic impairments that lie outside morphosyntax and reveal a deficit in general structural relations, such as difficulties with verb structure, noun phrases, WH-questions, embedding, and theta-role and pronominal/anaphoric reference assignment (Hamann, Penner & Lindner, 1998; Ingham, Fletcher, Schelleter & Sinkha, 1998; Jakobowicz, Nash, Rigaut & Gerard, 1998; Stavrakaki, 2001; van der Lely, 1998; van der Lely & Battell, 2003; van der Lely & Hennessey, 1999).

A hypothesis which is much more wide-ranging in terms of the phenomena it seeks to explain is van der Lely's Representational Deficit for Dependent Relations (RDDR) hypothesis (van der Lely, 1996, 1997b, 1998; van der Lely & Stollwerck, 1997). The RDDR proposes a deficit in the syntactic computational system. More specifically, the deficit resides in the inability to build up non-local relations between elements in a sentence, such as those required for WH-questions, binding relations, subject-verb agreement marking and embedding. As they are understood within the framework of the Minimalist Program

(Chomsky, 1995), these relations require 'Movement', which is driven by the need to check sublexical features. It is not that these movements or feature-checking are unavailable in SLI – just that their use is not obligatory. Although the linguistic explanation for the RDDR has evolved over the years, in the most recent model (van der Lely, 1998) the deficit is proposed to be in a principle of Economy termed 'Must Move'. As the name of this principle suggests, the result of the deficit is optional movement.[1]

Over the past three years, the direction taken by van der Lely, myself and our colleagues is that G-SLI is more accurately characterised as a deficit in structural complexity in three components of the grammatical system: syntax, morphology and phonology (and probably semantics as well, although this is yet to be investigated). This new position, the Deficit in Computational Grammatical Complexity Hypothesis, is an extension of the RDDR in that it retains the RDDR's explanation for the syntactic deficit, whilst acknowledging that morphological and phonological complexity cannot be defined in terms of structural non-local dependencies (van der Lely, 2004). Uncovering precisely how complexity in morphological and phonological representations is most accurately defined, and which aspects of this complexity most impact on typically developing and G-SLI children, is at the heart of the present thesis. In order to do so, an explicit theoretical linguistic framework is required, and this is set out in the next section with particular reference to morphology and phonology, together with a discussion of how typically developing and SLI children acquire morphology and phonology.

1.2. The theoretical framework

1.2.1. The components of language

Insights into the underlying nature of language disorders are only possible when empirical observations are analysed within an explicit theoretical framework. It is therefore essential that any model of language development in SLI be firmly rooted in linguistic and cognitive theory. Three components of language can be identified – the lexicon (i.e. the language user's mental dictionary), the computational system (in the sense of Chomsky, 1993) and the pragmatic system. The computational system is in turn composed of syntax, semantics, morphology and phonology – termed 'grammar' – and the mechanism for processing this information – the 'parser'. On this account, an impairment in language ability could affect one, or more or even all of the lexicon, syntax, semantics, morphology,

[1] Although the RDDR is currently framed in terms of the Minimalist Program, it is not tied to this framework – what is important is the characterisation of the deficit as being in the establishment of non-local syntactic relations.

phonology and pragmatics. Given the heterogeneity of SLI (see Section 1.1.2), it is likely that deficits in each of these areas are attested.

The modularity of the language system is a well-accepted construct in theoretical linguistics. Based on the work of Fodor (1983), modules are input systems with several essential properties: domain-specificity, informational encapsulation, mandatoriness, speed of operations, neural localisation and susceptibility to characteristic breakdown. Of those, domain-specificity is the most important characteristic of modules (see also Coltheart, 1999). Although the concept of modularity is frequently evoked in developmental studies, its psychological reality is not unanimously accepted by cognitive scientists (Levy & Kave, 1999). Even amongst those who do accept a degree of modularisation in the adult brain, there is debate over whether modules are innately specified or whether they emerge during development (Elman *et al.*, 1996; van der Lely, 1997b). It is useful to distinguish between 'big' modularity, which holds that language is a modular cognitive domain, and 'little' modularity, which refers to the internal, modular, organisation of language itself (Levy & Kave, 1999). Research into clinical groups that exhibit differential impairment and sparing of language versus general cognitive systems inform us about 'big' modularity. Studying differential impairment and sparing of aspects of language such as syntax and phonology in different subgroups of SLI can inform us about 'little' modularity. It is particularly the latter concept, that of 'little modularity', that the work in this thesis addresses.

Another fundamental notion that I adopt is Chomsky's concept of a Universal Grammar (UG; for a discussion, see Smith, 1999). UG ascribes a large part of the knowledge that speakers have about their native language to innate knowledge. UG consists of (1) a set of principles that are universal to all languages, and (2) a narrow set of parameterised choices that are language-specific, and whose values need to be learnt by the child during the process of acquisition (Chomsky, 1981). As we understand more about the genetic deficits underlying SLI, it may be possible to start relating patterns of linguistic breakdown to deficits in particular genes, and hence to characterise the genetic component of UG. Hence, it should be clear from the preceding discussion that just as characterising the linguistic deficit in SLI requires a theoretical linguistic framework, so studies of SLI can inform linguistic theory.

1.2.2. Morphological theory

Morphology comprises the following processes (Spencer, 1991):-

- **inflection** - the marking of syntactic relations, e.g. past tense marking as in *talk* → *talked*, aspectual marking as in *play* → *playing*,
- **derivation** - the changing of lexical category to yield a new word, e.g. adjective formation from nouns as in *spot* → *spotty*, agentive formation from verbs as in *bake* → *baker*,
- **compounding** - the joining of two or more words to make a new word, e.g. *green* + *house* → *greenhouse*.

The term 'morphosyntax' refers to how the form of a morpheme reflects its syntactic function, while 'morphophonology' refers to how the sound of a morpheme alternates in different phonological environments.

The status of morphology within linguistic theory is uncertain. Morphemes are at the interface of the physical form of language (its sound, i.e. phonology and phonetics) and its content (its meaning and function, i.e. semantics and syntax), but beyond that there are many views as to its relationship with other components of the language faculty. In early models of generative grammar, inflection, derivation and compounding were handled by the rules of syntax, with allomorphic variation (e.g. the realisation of the past tense affix as /t/, /d/ or /ɪd/ depending on the final consonant of the verb stem) regarded as the result of the operation of phonological rules (Chomsky, 1965). A range of models is currently proposed and defended – that morphology is located in the lexicon (di Sciullo & Williams, 1987; Lieber, 1992), located in syntax (Baker, 1985), distributed between syntax and the lexicon (Chomsky, 1995), distributed between various syntactic components (Halle & Marantz, 1993) or is a separate module independent of the two (Ackema & Neeleman, 2000; Froud, 2001).

The debate over the status of morphology as a whole is related to another issue that remains unresolved, namely the distinction between inflectional and derivational morphology. This distinction has proved impossible to define precisely. In general terms, however, inflectional morphology is considered to be productive, to leave the category of a word unchanged and to confer a regular meaning. Derivational morphology is considered more likely to be semi-productive, category-changing and idiosyncratic in meaning. In addition, derivational affixes tend to be closer to the root than inflectional affixes. Cross-linguistic observation shows that there are exceptions to all the aforementioned properties (Bauer, 1988), but despite difficulties of definition, most morphologists choose to keep the distinction (e.g. Anderson, 1992). It is generally assumed that inflectional morphology is accessible to, and therefore manipulable by, syntax whereas derivational morphology is not. For example, in the Minimalist Program derivational morphology is assumed to take place in the lexicon and be irrelevant to syntax (Chomsky, 1995).

A further unresolved debate concerns the mechanisms that underlie morphology, and in particular those that underlie regular morphology. Two particular models have a long history: one claims that regular suffixation is achieved by analogy, and the other claims that it is rule-governed. More specifically, the choice is between children learning to make the past tense form *bump* → *bumped* by analogy with the phonologically similar form *jump* → *jumped*, and between children using the abstract rule 'add *-ed* to all verb stems to express the past tense'.

The dual mechanism model, sometimes called the 'Words and Rules' (WR) model (Pinker, 1991, 1999; Ullman, 1999, 2001; Pinker & Ullman, 2002), proposes that different cognitive mechanisms underlie regular and irregular morphology:-

- Regular forms are generated by computational rules. These rules take affixes stored in the lexicon and combine them with stems, also stored in the lexicon (Halle & Marantz, 1993). However, the combination of a particular affix and a particular stem is blocked whenever an irregular form is retrieved from the lexicon. For example, *dug* blocks *dig+ed*.
- Irregular forms are stored in and retrieved from the lexicon by associative memory. Because the relation between irregular stems and their past tense is by definition arbitrary, each irregular stem has to be stored with information regarding its past tense form. This associative memory is partially productive in the sense that it will allow the generation of new irregulars for nonsense words that share certain phonological characteristics with existing irregulars, such as **spling* → **splang*, which follows the pattern of *sing* → *sang*, *ring* → *rang* and *spring* → *sprang*.

The two mechanisms underlying the regular-irregular distinction are claimed to be epiphenomena of the design of the human language faculty and are the same mechanisms that underlie the distinction traditionally made in linguistics between grammar and the lexicon.

In contrast to the WR model, the single mechanism connectionist model (e.g. Joanisse & Seidenberg, 1999; McClelland & Patterson, 2002; Plunkett & Marchman, 1991; Rumelhart & McClelland, 1986) claims that all verbs, both regular and irregular, are processed by just one mechanism – pattern association (i.e. analogy), and that rule-like behaviour can emerge from associational learning. This type of connectionist model is a pattern-associator network that learns to associate phonological features of the stem with phonological features of the past-tense form. The work in this thesis contributes to the debates on the single versus dual mechanism models of morphology, and on the distinction between inflectional and derivational morphology.

1.2.3. Typical acquisition of morphology

The wug test, which is passed successfully by children as young as four, shows that even young children have a mental rule for inflecting novel words (Berko, 1958). The wug test proves the creativity of morphology – no child could possibly have had the opportunity to memorise the plural of a nonsense word such as *wug*. English is not a completely regular language, however, and there are words which do not take the regular *-s* ending – words such as *mouse*, *goose*, *child*, *ox* and *deer*. The plurals of these words have to be memorised and stored in the lexicon.

The acquisition of inflectional morphology by young children proceeds through the following stages (I use the example of English plural *-s*):-

- All inflected plural words are memorised as such and stored as wholes in the lexicon.
- The child deduces the rule of inflection, i.e. 'add an *-s*', and starts to use it productively, creating forms such as *dogs* and *boys*, and over-productively, creating forms such as **mouses* and **childs*.
- As the associations between irregular words and their inflected forms become stronger in the lexicon, over-regularisations are blocked, so that the correct forms of irregular plurals are now given.

This sequence of acquisition gives rise to a U-shaped curve of correct forms plotted against time – there is a dip in performance as the rule is used over-productively and memory traces are not strong enough to block its application (Marcus, Pinker, Ullman, Hollander, Rosen & Xu, 1992).

The acquisition of compounding and derivation starts after children have started to acquire inflectional morphology (Clark, 1998). In English, the first novel word constructions emerge at around two years of age, allowing children to fill gaps in their lexicon (Clark, 2003). The first novel words are compounds, and verbs formed from nouns by zero-affixation (i.e. no audible suffix is added). At around age three, an increasing number of novel forms with affixes are produced; agentive and instrumental *-er* are two of the first to be used. There is evidence that children are aware from a young age of well-formedness constraints on word-formation. For example, children as young as three rarely include regular plural forms inside compounds (e.g. **rats-eater*), but they do produce compounds containing irregular nouns, e.g. *mice-eater*. Hence they show awareness of the more general rule that disallows inflection inside compounding and derivation (Gordon, 1985).

So when children create novel words, they use word-formation options that are legal options in the language they are acquiring, and they use well-established patterns that are productive in adult speech (Clark, 2003).

1.2.4. Acquisition of morphology in SLI

Inflectional morphology has been the most studied area of morphology in SLI because it causes difficulties that appear to be characteristic of the impairment. Focusing on this particular area in SLI research is valid because inflectional morphology has been extensively studied in the disciplines of theoretical linguistics, psychology, neurology, neuro-imaging and neural-network modelling (for a review, see Pinker, 1999). Therefore we have a good theoretical understanding of how the development of inflectional morphology proceeds in normally-developing individuals, enabling researchers to speculate as to the underlying cause of the deficits observed in SLI.

It is well-documented that the majority of English-speaking children with SLI face immense difficulty with inflectional morphology (see Leonard, 1998, for a thorough review). They use inflection only optionally in contexts where it is obligatory, even for the same lexical item (e.g. Rice *et al.*, 1995; van der Lely, 1997b). For example, in a narrative test based on the book 'Frog, Where Are You?', one child used the forms *fell* and *fall* interchangeably in a past tense context, while another used both *looked* and *look* and a third *came* and *come* (van der Lely, 1997b and unpublished data; see also Leonard, Eyer, Bedore & Grela, 1997). Indeed, difficulties with inflection are so widespread that Rice and Wexler (1996) have proposed that tense-marking be used as a clinical marker for SLI. However, when inflection is used, it is used correctly, and only very rarely misapplied to stems in contexts that are not allowed in adult speech (Rice & Wexler, 1996; Miller & Leonard, 1998). Errors such as **They sings* and **They's shut* are almost non-existent. SLI children are capable, as typically developing children are, of using inflection productively, making irregularisations such as **throwed* for *threw* (e.g. Marchman, Wulfeck & Weismer, 1999; Oetting & Horohov, 1997).

Dual mechanism models (see Section 1.2.2) predict particular patterns of morphological breakdown – either rule application may be impaired, or blocking may be impaired. SLI is a likely candidate for the former (Pinker, 1991), and the theory is claimed to explain data in studies of both plural-formation (Goad, 1994) and past tense morphology (van der Lely & Ullman, 2001). Van der Lely and Ullman looked at past tense production of both real and novel regular and irregular words in a group of G-SLI children and three groups of language-age matched children. They found two patterns of interest in the data:

- The control children showed a significant advantage for regular verbs over irregular verbs, whereas the G-SLI children showed no such advantage.
- For the G-SLI group but not the control groups, successful production of regular past tense forms improved as a function of frequency,

Van der Lely and Ullman interpreted their findings within the framework of the dual mechanism model. The lack of regularity advantage provides evidence of an impaired morphological rule which is, however, available to the typically developing children. G-SLI children have to rely instead on the preferential storage of past tense forms, hence the frequency effect found for regulars. Note too the persistence of their deficit: the G-SLI children in van der Lely's experiment were aged between 9;03 and 12;10, and correctly inflected only 22% of regular forms.

Single mechanism models of morphology are, however, argued to also be able to capture the pattern of impairment in SLI (e.g. Joanisse & Seidenberg, 1998; Joanisse, in press). Such models capture the dissociation between regular and irregular verbs in terms of damage to a single system. Speech input is mapped directly onto a set of distributed representations with a specific architecture of connections, modelling words in terms of their speech input, speech output and semantics. Lesioning the speech output layer models a phonological deficit, which affects past tense generation performance on non-words and regular past tense verbs. Lesioning the semantic layer models a deficit in irregular past tense formation. In Joanisse and Seidenberg's model there is no need for any explicit morphological differentiation between regular and irregular forms, because the differences between them reflect the relative balance between semantic and phonological factors during the acquisition phase of the network. In contrast to dual mechanism models, single mechanism models do morphology without the mediation of linguistic rules, and indeed, without any conventional linguistic representations.

1.2.3. Phonological theory[2]

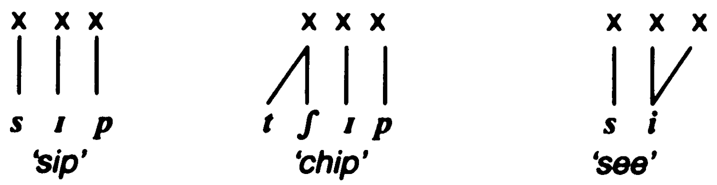
1.2.3.1. A model of prosodic complexity

A phonological representation comprises segmental information (the 'melodic' aspect) and prosodic information (the 'structural' aspect). Non-linear accounts of phonology (e.g. Goldsmith, 1976) allow the independent representation of segments and prosody, which are linked via the skeletal tier. Consonants, including affricates, take up only one slot of the

[2] There is confusion in the psycholinguistic literature surrounding the use of the terms 'phonology' and 'phonetics'. I reserve 'phonology' to cover only the organisation of speech sounds in the grammatical system, and phonetics to refer to the physical properties of those speech sounds.

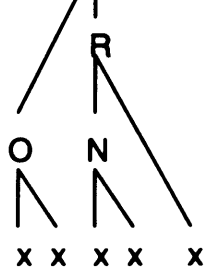
skeletal tier (represented by x in Figure 1.1), as do short vowels. Long vowels and diphthongs take up two slots. Segmental material is linked to the skeletal tier via 'association lines', as shown in Figure 1.1.

Figure 1.1. The linking of segments to the skeletal tier



Prosody defines relations within phonological strings, and prosodic complexity can be defined in terms of hierarchical structure (e.g. Hayes, 1980; Kager, 1989). The syllable is a fundamental unit of the prosodic hierarchy. According to the most widely accepted model of syllable structure, a syllable consists of an onset and a rhyme, with the rhyme in turn composed of a nucleus and a post-nuclear slot called the coda (see Blevins, 1995). However, Harris (1994) argues against there being a separate coda constituent. I adopt Harris' model of syllable structure, while recognising that he questions the existence of a separate syllable node, considering instead that it is a relation between the onset and rhyme. Because the notion of 'syllable' is so widely used in the theoretical and psycholinguistic literature I will not abandon it. The structure, illustrated in Figure 1.2, is binary branching – the onset (O), rhyme (R) and nucleus (N) may each have one or two segments linked to them, while segmental material can be missing altogether from the onset. The branching of syllabic constituents allows an obvious way of defining prosodic complexity – constituents that branch are complex and those that do not branch are simple.

Figure 1.2. Syllable (S) structure

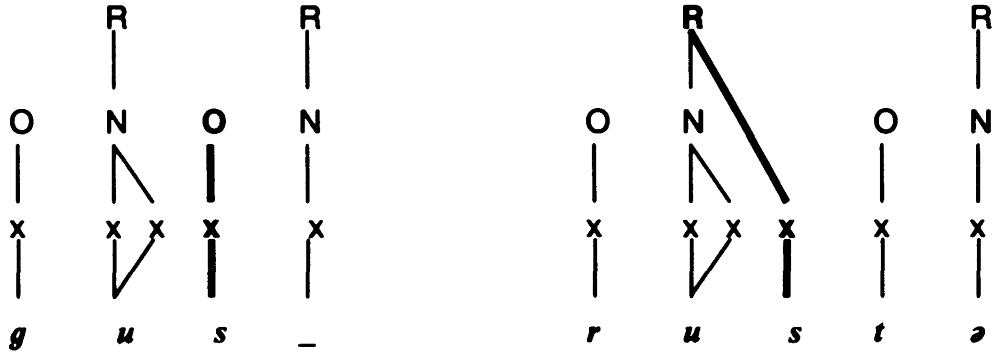


An ongoing debate in the phonological literature (e.g. Harris & Gussman, 2002; Piggott, 1999) concerns the status of word-final consonants. I follow Harris (1994) in categorising a word-final consonant as an onset to a syllable whose nucleus is empty. This means that, for example, the /s/ of *goose* and the /s/ of *rooster* occupy different positions within the syllable, as shown in bold in Figure 1.3. In traditional treatments of English phonology, the consonant of a branching rhyme and a word-final consonant are typically conflated, reflecting the assumption that a word-final consonant belongs to the same syllable as the preceding vowel. This assumption is in fact contradicted by a wide range of evidence, including syllable typology and consonant phonotactics (see discussion in Harris & Gussmann, 2002).

Figure 1.3. Different syllabic locations of a post-nuclear consonant

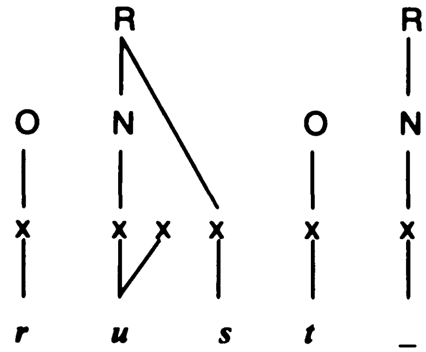
(i) Word-final

(ii) Word-internal



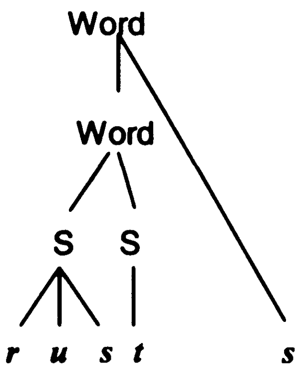
The number of consonants that can be supported word-finally in unsuffixed English words is limited to two. The syllabic structure of a word-final cluster is shown for the word *roost* in Figure 1.4. It can be seen that the two consonants occupy the same location as they did in the word *rooster*.

Figure 1.4. Syllabic structure of a word-final cluster



Inflectional morphology is most frequently signalled in English by the addition of a consonantal suffix to the right edge of the word (e.g. Goad, White & Steele, 2003). This structure is shown (with syllable structure simplified) in Figure 1.5 for the form *roosts.*

Figure 1.5. Prosodic structure of a consonantal suffix



We also need to consider complexity above the level of the syllable. Directly dominating the syllable is a unit called the foot, which can be viewed as supplying rhythmic organisation to the phonological string. The English foot conforms to several of the patterns associated with unmarked metrical structure: it is binary; it establishes the size of the minimal word; and it is trochaic, displaying a left-dominant stress pattern (see Hayes, 1995). Binarity is satisfied by two weight-bearing positions: the foot consists of either disyllabic trochee (as in *city*), or a monosyllable consisting of a long vowel (as in *tea*) or a vowel followed by a consonant (as in *sit*). The foot is the level at which stress is assigned. In English, stress is manifested phonetically by increased loudness, pitch and duration. In addition, vowel quality is reduced in unstressed syllables.

Above the level of the foot is a further level – the prosodic word, which must minimally contain one foot. The prosodic hierarchy can be schematised as in Figure 1.6.

Figure 1.6. The prosodic hierarchy

Prosodic word
|
Foot
|
Syllable
|
Skeletal tier

1.2.3.2. Optimality Theory

In this thesis I examine several phonological phenomena within an Optimality-Theoretic framework. Optimality Theory (OT) is a framework for transforming an underlying linguistic representation into a surface representation (Prince & Smolensky, 1993). For example, during acquisition a child might produce *blue* (the underlying, input form) as ***/bu/*** (the surface, output form). An OT grammar has a set of ranked constraints that choose the optimal output form from a set of candidate outputs. Constraints that are more active are said to be higher ranked than those that are less active. Constraints are universal and languages vary only in how they rank them. Under this view, when a child learns his language he learns the language-specific rankings of those constraints.

Constraints can be divided into two groups – faithfulness and markedness constraints. Faithfulness constraints regulate the correspondence between an output and its input, whereas markedness constraints regulate the structure of the output and place restrictions on that structure. It is generally agreed that in child phonology markedness constraints are initially ranked higher than faithfulness constraints. This means that pressure to produce simple, unmarked structures is stronger than the pressure to produce an output that is identical to the input.

For example, in English onsets may be simple (as in ***/bu/***) or complex (as in ***/blu/***). Let's continue with the example of a child simplifying ***/blu/*** to ***/bu/***. Two constraints are relevant to the onset in this case: a faithfulness constraint which ensures that the onset in the output is identical to the onset in the input, and a markedness constraint that does not allow complex onsets to appear in the output. For the sake of transparency, I'll call these FAITHONSET and NOCOMPLEXONSET respectively. The child is assumed to start off with a grammar that ranks the markedness constraint, NOCOMPLEXONSET, higher than the faithfulness constraint, FAITHONSET. This is shown in the tableau in Figure 1.7.

Figure 1.7. NOCOMPLEXONSET >> FAITHONSET

	/blu/	NOCOMPLEXONSET	FAITHONSET
a.	***blu***	*!	
b. ☞	***bu***		*

In OT tableaus, constraints are shown heading columns. The further left a constraint appears in a tableau, the higher ranked it is, and therefore the more active it is. Candidate forms are shown below the input form, which is assumed to be the child's underlying representation. In this example there are two candidates relevant to our discussion – ***/blu/***, where the onset is produced correctly, and ***/bu/*** where it is simplified. The asterisks show which constraints are violated by which candidate. NOCOMPLEXONSET is violated by ***/blu/***, because the output contains a complex onset, and FAITHONSET is violated by /bu/, because the output is not faithful to the underlying representation. However, because NOCOMPLEXONSET is ranked higher than FAITHONSET, it is the candidate that violates this constraint that 'loses' and therefore ***/bu/*** that 'wins'. The exclamation mark following the asterisk indicates that this violation is fatal, while the pointing hand indicates the winning candidate.

This same constraint ranking ensures that the word *boo* is indeed pronounced as ***/bu/***. *Boo* is not pronounced as ***/blu/***, not only because ***/blu/*** would violate NOCOMPLEXONSET, but also because it is not faithful to the onset in the underlying representation. This is shown in the tableau in Figure 1.8.

Figure 1.8. NOCOMPLEXONSET >> FAITHONSET

	/bu/	NOCOMPLEXONSET	FAITHONSET
a. ☞	***bu***		
b.	***blu***	*!	*

Important questions for OT are how to define markedness and how to determine what sorts of segments and prosodic structures are dispreferred. The respective properties of unmarked versus marked segments/structures are as follows (Blevins, 1995; Harris, 1994):-

- All languages have words with unmarked segments/structures, but only some have words with marked segments/ structures.
- If a language has words with a marked segment/structure, it will also have words with the unmarked segment/structure. The reverse implication is not true.
- In acquisition, unmarked segments/structures are acquired first.
- In second language acquisition, speakers have little difficulty in going from a marked to an unmarked segment/structure, but do show difficulty in going from the unmarked to the marked.

For example, the marked status of complex onsets is confirmed by the fact that many languages lack them altogether as well as by the fact that, in languages that do have them, they are acquired later than simplex onsets. Other marked structures include branching rhymes and word-final consonants. A theory of markedness allows us to make predictions about which prosodic structures children with SLI will find difficult.

1.2.3.3. Typical acquisition of phonology

Typically-developing children produce simple, unmarked syllabic structures before they produce more complex, marked ones (Bernhardt & Stemberger, 1998). For example, young children produce non-branching onsets before branching ones. Although it is well-known that there are limits on the types of syllables that children can initially produce, there has been little work on the acquisition of syllable structure.

The most comprehensive study is perhaps that by Levelt, Schiller and Levelt (2000) on Dutch children. They posit that the acquisition of syllable types follows the sequence presented in Figure 1.9.

Figure 1.9. Stages in the acquisition of syllable structure (Levelt *et al.*, 2000)

$$\mathrm{CV} \rightarrow \mathrm{CVC} \rightarrow \mathrm{V} \rightarrow \mathrm{VC} \begin{matrix} \nearrow \mathrm{CVCC} \rightarrow \mathrm{VCC} \rightarrow \mathrm{CCV} \rightarrow \mathrm{CCVC} \searrow \\ \searrow \mathrm{CCV} \rightarrow \mathrm{CCVC} \rightarrow \mathrm{CVCC} \rightarrow \mathrm{VCC} \nearrow \end{matrix} \mathrm{CCVCC}$$

In this analysis, V and VV are both treated as V, as it seems that vowel length is not initially contrastive (Fikkert, 1994). A note of caution is that these syllable types are based on single words, and Levelt *et al.* see all word-final consonants as codas, in contrast to the view of syllable structure that I adopt in Section 1.2.3.1. The only variation in this sequence is that some children (9/12) acquired word-final clusters before complex onsets, while others (3/12) acquired them in the opposite order. Indeed, there is no theoretical reason why one structure should emerge before the other. Note that there is a point during development when both complex onsets and complex word-final clusters are allowed, but not within the same syllable.

Although Levelt et *al.*'s work is on Dutch, work on the acquisition of English syllable structure supports their model. For example, Bernhardt and Stemberger (1998) report similar stages, although they stress that there may be considerable variation from child to child (Stemberger & Bernhardt, 1999). There is also variability with regards to the order of acquisition of particular combinations of segments within clusters (Bernhardt & Stemberger, 1998). Kirk and Demuth (2003) report that word-final clusters are acquired before onset clusters, with the exception of /s/ + stop clusters.

As for the acquisition of foot structure, Echols and Newport (1992) note that English-speaking children tend to include stressed syllables and final syllables in their early speech, but that they omit non-final unstressed syllables. The classic error of this type is the word *elephant* being reduced to */efənt/* by omission of the middle unstressed syllable. The tendency to delete word-initial weak syllables is also well-noted (Bernhardt & Stemberger, 1998; Fikkert, 1994), and is due to constraints on the grammar rather than any perceptual difficulty (Kirk & Seidl, 2003).

Gerken (1994) observes that children's early utterances tend to be organised into strong-weak feet. This means that children find it easier to produce an unstressed syllable which follows rather than precedes a strong foot. For example, in a sentence such as 'the DOG + KISSED her', where the article forms a prosodic word with the following noun, the article is omitted more frequently than in a sentence such as PETE + KISSED the + DOG, where it can form a prosodic word with the preceding verb (Gerken, 1991). Gerken uses the analogy of a template for foot structure. The child has at his or her disposal initially just one, but then increasingly more templates. Children will modify words and phrases to fit this/these template/s. As more templates become available, their language gradually becomes more target-like.

The frequency with which particular structures occur in the input is known to affect the rate at which children acquire marked prosodic structures. For example, initial weak syllables are acquired earlier in Spanish than in English, while word-final consonants are acquired earlier in English than in Spanish; this correlates with initial weak syllables being more frequent in Spanish and word-final consonants being more frequent in English (Roark & Demuth, 2000).

1.2.3.4. Acquisition of phonology in SLI

In contrast to inflectional morphology, the status of phonology in SLI is still poorly understood (Leonard, 1998), although it is recognised as being deficient in many children (Bortolini & Leonard, 2000). It is also recognised that any full account of SLI will need to incorporate phonology (Bortolini, Caselli & Leonard, 1997).

Early studies on segmental production showed, perhaps unsurprisingly, that children with SLI are late in acquiring the segments of a language (for a review, see Leonard, 1998). Other studies provided evidence that individuals with SLI show a tendency to produce open (i.e. CV) rather than closed (i.e. CVC) syllables, and syllables with simple rather than branching onsets (Schwartz, Leonard, Folger & Wilcox, 1980; Ingram, 1981). More recently, Bortolini and Leonard (2000) showed that English-speaking SLI children omit significantly more consonants in word-final position than normally developing controls matched on mean length of utterance, and reduce the majority of word-final consonant clusters.

Work on Italian shows that SLI children have difficulty in representing complex syllabic structures (Orsolini, Sechi, Maronato, Bonvino & Corcelli, 2001). In disyllabic words there is a strong tendency to simplify the onset of the first syllable if it is complex (e.g. *grande* becomes *gande*). Yet those children don't simplify the rhyme of the first syllable from CVC to CV. If that post-nuclear consonant is simplified, it becomes geminate with the onset of the second syllable (e.g. *por.ta* is simplified to *pot.ta* rather than to *po.ta*, where . indicates the syllable boundary).

Impairments of metrical structure are also evident in SLI. Sahlen, Reuterskiold-Wagner, Nettelbladt and Radeborg (1999) found that Swedish-speaking children with SLI aged 4;11-5;11 omit unstressed syllables six times more often in weak-strong positions than in strong-weak positions in both words and non-words, although the proportion of omitted syllables in weak-strong position only amounted to 7% of the total number of target weak syllables in that position. Sahlen *et al.* provide no data comparing the performance of their SLI subjects to typically-developing controls, so it is impossible to infer that prosody is affected to a greater extent in SLI than would be expected at that stage of language development. However, evidence from Dutch suggests that SLI children (aged 3;04-4;07) do indeed omit initial weak syllables significantly more frequently than younger, typically developing children in both real and non-words (de Bree & van der Pas, 2003).

Carter and Gerken (2002) have investigated initial weak syllable omissions in English-speaking children with SLI, aged 4;02-6;01. These children produced sentences containing reduced or unreduced disyllabic proper names (e.g. 'Feed _cinda' from 'Feed Lucinda', 'Feed Cindy'). The duration of verb-onset to name-onset in both types of sentences was measured, and found to be significantly longer in those containing reduced names. This indicates that even when the first syllable of a name such as *Lucinda* was omitted an acoustic trace remained. These results can be interpreted as indicating that although the segmental material was omitted, the timing slot remained.

There is evidence that segmental accuracy is affected by position within the syllable and/or the foot, in both typically developing (e.g. Inkelas & Rose, 2003) and language-impaired children. For example, Marshall and Chiat (2003) report on a child with a phonological disorder who replaces fricatives (e.g. */s/*) with stops (e.g */d/*) in foot-initial positions, but who realises fricatives correctly when they are foot internal or foot-final, while Marshall, Ebbels, Harris and van der Lely (2002) report on a child who replaces */ʃ/* with */θ/* foot internally. Interactions between syllabic complexity and metrical complexity, however, have not been explored in either typically developing or language-impaired children.

1.3. Morphology and phonology interactions in typical development and SLI

1.3.1. How morphology and phonology interact in typical development

Inflection and derivation in English typically take place through suffixation of a segment, e.g. *play* → *plays*, *yell* → *yelled*, or a whole unstressed syllable, e.g. *horse* → *horses*, *whisper* → *whispering*, *big* → *biggest*, *sand* → *sandy*. Adding a suffix that consists of a single segment changes the syllabic structure of the word end. This means that a word whose root form ends in a vowel now ends in a single consonant, e.g. *plays*, while a word whose root form ends in a single consonant now ends in a cluster, e.g. *yelled*. Adding a syllabic suffix alters the metrical structure of the word, with the unstressed syllable sometimes falling within the foot, e.g. {*bigger*}, {*sandy*}, but sometimes outside the foot, e.g. {*whispe*}*ring* (where brackets indicate the foot boundary).

Given that word-final consonants, word-final clusters and unfooted syllables are all marked prosodic structures (see Section 1.2.3.2), it is predicted that phonological constraints will impact on the realisation of inflection (Bortolini & Leonard, 2000; Stemberger & Bernhardt, 1997). In other words, a child who is unable to produce word-final consonants, word-final clusters and unfooted final syllables in monomorphemic words is not expected to produce them in morphologically complex words. In support of this prediction, Stemberger and Bernhardt (1997) report a child who reduces clusters in both monomorphemic and suffixed words, e.g. *fox* → /*fɔk*/ and *rocks* → /*wɔk*/. However, the pattern is not always so straightforward. Stemberger and Bernhardt report a different child who produces word-final clusters *only in suffixed forms*. They also report data from a child who reduces monomorphemic clusters differently to clusters in suffixed forms, e.g. *fox* → /*fɒs*/ but *rocks* → /*wɒt*/ (c.f. *rock* → /*wɒt*/). Therefore, the way in which children treat complex structures in monomorphemic forms does not necessarily match how they treat them in suffixed forms.

Not only does inflection introduce marked prosodic structure at the word end, it also, on occasion, introduces sequences of segments that do not occur in monomorphemic words. In other words, inflection can introduce sequences that are phonotactically illegal within morphemes, for example the /*md*/ of *slammed* or the /*gd*/ of *hugged*. These sequences are frequently marked cross-linguistically, meaning that only certain languages will allow them to occur monomorphemically. These sequences that are illegal in English are therefore more 'difficult' in some way compared to clusters such as /*nd*/ and /*kt*/ which English does allow to occur monomorphemically. There are therefore two ways in which inflection introduces phonological difficulty at the word end: it introduces prosodically marked structure and it introduces phonotactically marked sequences of segments.

1.3.2. Does a phonological deficit cause the morphological deficit in SLI?

Some accounts of SLI claim a causal link between deficits in phonology and difficulties in morphosyntax. Those who assume a domain-general perspective contend that lower-level input-processing deficits cause all forms of SLI. Here, phonological deficits are considered to be at the interface of the language deficits and either defective auditory processing, short-term memory, or limited capacity or speed of processing. For example, slow processing is hypothesized to cause problems with processing sounds with rapid acoustic transitions (/*t*/, /*d*/) and/or perceiving phonemes with low-phonetic salience (/*t*/, /*d*/, /*s*/ and /*z*/). This, in turn, causes problems with certain inflections such as the past tense allomorphs /*t*/ and /*d*/ (*jumped, played*) and agreement/tense and plural /*s*/ and /*z*/ (*jumps, pens*) (Joanisse & Seidenberg, 1998; Leonard, 1998).

On the other hand, accounts within the linguistic perspective have failed to consider the effects of phonology. For instance, Bernhardt and Stemberger (1998) dispute Gopnik's (1990), Gopnik and Crago's (1991) and Pinker's (1991) conclusions that inflectional morphology is impaired in the KE family, and argue that phonological factors were discounted. The basis of Bernhardt and Stemberger's argument is that outlined in Section 1.3.1, namely that if clusters are not allowed word-finally in monomorphemic forms, then it is reasonable to expect that they will not occur in inflected forms (Bernhardt & Stemberger, 1998). Similarly, Stemberger (personal communication) has criticised van der Lely and Ullman's (2001) study of past tense morphology by again arguing that phonological factors were ignored. Connectionist models show that when phonology is taken into account, the severity of the morphosyntactic impairment is predictable (Stemberger, 1995). However, Stemberger's claim for a phonological deficit rather than a morphosyntactic one is in turn too simplistic. Van der Lely has never denied that phonology could indeed be impaired in G-SLI, and has never discounted the fact that there may be independent deficits in syntax and phonology which interact in inflectional morphology (e.g. van der Lely, 1997b, 1998). Importantly, it remains to be established whether all children with morphological deficits suffer phonological deficits.

Bortolini and Leonard (2000) addressed the link between a phonological deficit and a morphological deficit by looking at SLI children's rates of final consonant deletion and final cluster reduction in monomorphemic words, and comparing them with rates of inflection omission. Children with the highest rates of final cluster reduction were the ones who omitted inflections most frequently. The authors did not find a correlation between final consonant deletion and inflection omission, even though the SLI children deleted final consonants more frequently than their language-matched controls. This may be because any child who omitted more than 20% final */s/*, */z/*, */t/* and */d/* from monomorphemic words was excluded from the study, while final consonant deletion was measured over all consonants, not only those used for inflection. Nor did the authors have a separate measure for inflection omissions where the suffix is part of a cluster as opposed to where it is a singleton consonant. Ideally we would want these two separate measures to compare with phonological behaviour on monomorphemic words.

Few studies of inflectional morphology in SLI have taken into account the phonology of the inflected verb-ending. Of those that have, one has concluded that phonology has no effect on SLI performance. Oetting and Horohov (1997) examined past tense marking in a group of 6-year olds with SLI and found that they, in common with their language-matched controls, inflected verbs ending in vowels, glides and liquids more frequently than those ending in a fricative or stop. Even if phonology does indeed affect both groups similarly, Oetting and Horohov do not link their claim with any theory of phonology that could account for why verbs ending vowels, glides and liquids should be easier for both groups of children. They therefore miss the generalisation that the word-end of an inflected verb whose stem ends in a vowel or glide does not have a consonant cluster when inflection is added, and is therefore has a simpler prosodic structure than a verb stem that ends in a consonant and therefore *does* have a cluster when inflected. Next, it appears that while control children do not treat verbs ending in a consonant differently according to whether the verb takes */t/* or */d/*, the SLI children do – they perform worse on those taking */d/* (although this difference is not tested for significance). Oetting and Horohov miss the generalisation that a cluster ending in */d/* is more likely to be phonotactically illegal than one ending in */t/*. For example, only */nd/* and */ld/* (i.e. sonorant that agrees in place of articulation + */d/*) are legal word-final clusters, whereas */bd/*, */gd/*, */dʒd/*, */md/*, */vd/*, */zd/* and */ðd/* (i.e. voiced obstruent, or sonorant that does not agree in place of articulation + */d/*) are all illegal. On the other hand, */kt/*, */ft/*, */pt/*, and */st/* are all legal while only */ʃt/* and */tʃt/* are illegal. It is possible that SLI children are sensitive to legality whereas their typically developing matches are not.

The impact of phonology on past tense inflection in G-SLI has not yet been investigated. However, in a case study of one particular G-SLI boy, AZ, van der Lely (1997b) noted that the only verbs for which AZ produced inflection in her and Ullman's elicitation task were *mar* and *stir*; those past tense forms, *marred* and *stirred*, contain sequences which can occur as part of the stem of other words, e.g. *hard/card* and *word/ bird*. AZ did not inflect regular verbs where the addition of the inflection would produce a sequence not attested in the stem form of words, e.g. *slammed*, *rushed*. This observation suggests that phonological factors at the verb-end may indeed play a role in past tense morphology in the G-SLI subgroup, although it is unclear from these examples what the relevant factor is. The difficulty could lie with segmental factors (e.g. the phonotactic sequences of consonant clusters) or prosodic factors (e.g. limits on syllable constituent branching or indirect licensing of extrasyllabic elements).

Critical to the debate as to whether phonology impacts on inflection in SLI is the difference in phonology between regular and irregular forms. Recall that SLI children have a deficit in regular versus irregular past tense formation relative to typically developing children. Irregular past tense forms always contain sequences that are found in well-formed single morphemes; indeed, many have monomorphemic homophones (e.g. *made/maid*; *sold/hold*; *wrote/boat*). Regular morphology, as has already been discussed, often (but not always) introduces marked phonology. This implies that if regular inflectional morphology breaks down, this is the result either of a difficulty with regular (as opposed to irregular) inflection or of a difficulty with marked phonology, or both. Teasing these two apart is not trivial, and is a major aim of this thesis.

1.4. Structure of the thesis

The thesis is structured as follows. Chapter 2 completes "Part 1. The background". I present a characterisation of the G-SLI subgroup and details of the children who participated in my studies. The experimental chapters are divided into 3 parts. In Part 2 I establish that G-SLI children have independent deficits in morphology (Chapters 3 and 4) and phonology (Chapter 5). In Part 3 I characterise in detail how phonology impacts on past tense inflection in typically developing children and those with G-SLI. Chapter 6 investigates the impact of verb end clusters on past tense judgements of regular and irregular verbs. Chapter 7 investigates the impact of verb-end cluster complexity on the production of regular past tense forms. Chapter 8 investigates the use of the syllabic past tense allomorph (i.e. */ɪd/*) with regular and irregular verbs. I start to outline a model of linguistic impairments in G-SLI called the deficit in Computational Grammatical Complexity (CGC) hypothesis, whereby independent deficits in syntax, morphology and phonology impact on past tense formation. In Part 4 I characterise how phonology impacts on less well-studied aspects of morphology – plural and progressive inflection (Chapter 9), and derivation (Chapters 10 and 11). I argue that the findings from these studies are consistent with the CGC hypothesis. In the final chapter (12) I summarise my research findings, discuss what future research is needed for the development of the CGC hypothesis and what issues my thesis raises for linguistic and cognitive theory.

As a note of guidance to the reader: I present my experimental results in tables, and only illustrate the data using figures when there is an interaction between variables.

Chapter 2. Grammatical-SLI

2.1. Introduction

2.1.1. Chapter outline

Given the linguistic heterogeneity of SLI, it is theoretically undesirable to generalise experimental findings to *all* children with SLI. Ideally, findings should be generalised to a small group of children who share the same, well-defined language characteristics. Nevertheless, the existence of Grammatical-SLI, claimed by van der Lely and her colleagues to be a homogeneous subgroup of the SLI population, is controversial. The aims of this chapter are to both define the characteristics of the G-SLI children who participate in my studies and to explain the theoretical reasons for studying this group. I have already introduced the linguistic and non-linguistic characteristics of G-SLI in Section 1.1.2, but I recap them here in Section 2.1.2, and introduce in Section 2.1.3 the Representational Deficit for Dependent Relations (RDDR) hypothesis that has been proposed to account for the syntactic deficit. In Section 2.2 I discuss how participants are selected to the G-SLI subgroup. I then discuss the selection of typically developing control groups (Section 2.3), and the controversy over whether G-SLI exists as a homogeneous subgroup (Section 2.4).

2.1.2. Linguistic and cognitive characteristics of G-SLI

G-SLI is a language deficit that is confined to the core aspects of grammar – syntax, morphology and phonology – and that persists beyond the age of 8 years. Non-local syntactic dependencies such as subordinate clauses (van der Lely, Rosen & McClelland, 1998), WH-questions (van der Lely & Battell, 2003), reversible passives (van der Lely, 1996) and pronominal and anaphoric reference (van der Lely & Stollwerck, 1997) are all impaired, in both production and comprehension. Morphology is likewise impaired, with children showing high omission rates of past tense inflection (van der Lely & Ullman, 2001), and producing regular plurals inside compounds (van der Lely & Christian, 2000). They also suffer subtle deficits in prosody that affect syllabic and metrical structure (Gallon, 2002; Marshall, Harris & van der Lely, 2003; Peiris, 2000). However, all have non-verbal IQ scores in the normal or above-normal range, and the group as a whole does not evince consistent auditory deficits (van der Lely, Rosen & Adlard, in press).

Van der Lely claims that 10-20% of children with persistent SLI and non-verbal IQs above 85 have G-SLI (van der Lely & Stollwerck, 1996), which translates into an incidence of approximately 3/1000 (van der Lely, personal communication, June 2003). Preliminary investigations of the family histories of some of these children reveal a pattern of impairment consistent with an autosomal dominant inheritance (van der Lely & Stollwerck, 1996), and further work is planned to identify the locus of the putative gene(s). Note that this does not imply that grammar is controlled by a single gene, but rather that a genetic deficit can have an impact on one or more of the mechanisms underlying grammar.

2.1.3. The Representational Deficit for Dependent Relations (RDDR) hypothesis

The RDDR hypothesis aims to account for the broad range of syntactic errors made by the G-SLI group (van der Lely, 1998; van der Lely & Battell, 2003). It locates the underlying deficit in the computational syntactic system, and uses Chomsky's Minimalist Program (Chomsky, 1993, 1995) as a framework for describing and explaining this deficit. Uniting the range of syntactic deficits is their reliance on the construction of non-local structural dependencies. For example, interpretation of 'Baloo Bear says Mowgli is tickling him' requires a dependency relation between *him* and the non-local antecedent *Baloo Bear*. In the Minimalist Program a dependent structural relation is formed between constituents in a sentence for the purpose of linking and checking grammatical features associated with lexical items: only by such feature checking can these items be given an interpretation. For example, formation of the object WH-question 'Who did Ralf see (t_i)?' requires a non-local dependency between the WH-operator (*who*) in the specifier position of the complementizer phrase and the trace in the internal verb argument position (t_i) which is bound by the operator. The WH feature on *who* is checked in the sentence through 'Movement' (Chomsky, 1995), and van der Lely claims that Movement, which is obligatory in normal grammar, is only optional in G-SLI grammar. It is not that the Movement operation is missing, or that the features that trigger movement are missing, but rather that the syntactic operation that forces Movement (termed 'Must Move') is impaired (van der Lely, 1998). Crucially, the RDDR hypothesis predicts that not all aspects of syntax will be impaired, just those that rely on non-local dependencies. So negation, for example, in which the negative particle is merged into the numeration, is not predicted to be impaired, and indeed it is not: when producing simple negative sentences such as 'They aren't on their skateboards', G-SLI children never once omit the negative particle *not* or *n't* (Davies, 2001).

If non-local dependencies are not reliably built in G-SLI grammar, then particular errors are predicted. When dependencies are built, comprehension/production will be correct, and when they are not built, comprehension/production will be incorrect. This will result in optionality in children's use of certain syntactic structures when non-syntactic cues are unavailable. For example, 'the boy is chased by the dog' will be interpreted either as the 'the dog chases the boy' or 'the boy chases the dog'. However, when unambiguous semantic and world knowledge cues are available, for example 'the apple is eaten by the man', children should be able to use these cues to interpret the sentence correctly – men can eat apples but apples cannot eat men. Similarly, the interpretation of 'Baloo Bear says Mowgli is tickling him' relies on syntactic knowledge for the interpretation of *him*'s antecedent, which is ambiguous between *Baloo Bear* and *Mowgli*. 'Baloo Bear says Cinderella is tickling him', however, is unambiguous, as the only available male antecedent is *Baloo Bear*.

The RDDR can also account for difficulties in past tense marking which are found for both regular and irregular past tense verbs, and in difficulties with third person agreement. Movement for the checking of tense and agreement features needs to take place between the verb and the inflectional phrase, and if this does not take place, then inflection will not be realised. The picture for morphology is further complicated by the finding that regularly inflected forms show a greater deficit in comparison to irregulars than would be predicted if G-SLI children were showing delayed but normal development (van der Lely & Ullman, 2001). Van der Lely and Ullman claim that G-SLI children suffer from a deficit in morphological rule use.

It will become clear from the work presented in this thesis that not only do G-SLI children suffer syntactic and morphological deficits, but that their phonology is also affected. Hence the RDDR needs to be revised in order to account for the broader range of impairments that affect three components of grammar – syntax, morphology and phonology. A new model, the Computational Grammatical Complexity (CGC) hypothesis has been developed over the last couple of years by van der Lely, myself and our colleagues, and I will return to this model in more detail in the concluding chapter of this thesis.

2.2. Selection of G-SLI participants

Selecting G-SLI participants is a two-stage process. In the first stage, children between the ages of 8 and 16 who have received a diagnosis of SLI are recruited from residential language schools or from language units within day schools. This recruitment is done with the help of speech and language therapists, who are asked to select only children with normal hearing and articulation, with English as a first language, and without a diagnosis of autistic spectrum disorder. Non-verbal intelligence tests are administered (e.g. British

Ability Scales, BAS, Elliot, 1996; Raven's Progressive Matrices, RPM, Raven, 1998) to ensure that we only select children with non-verbal IQ scores of greater than one standard deviation below the mean (i.e. a standard score greater than 85). Scores from standardised language tests, including the Test for Reception of Grammar (TROG; Bishop, 1983), British Picture Vocabulary Scales (BPVS; Dunn, Dunn, Whetton & Burley, 1997) Test of Word-Finding (TWF; German, 2000) and Clinical Evaluation of Language Fundamentals (CELF; Semel, Wiig & Secord, 1995) are obtained, often from the child's speech and language therapist, in order to build up a profile of the child's general language abilities. Children who have been recruited in this way, and who show a pattern of a more severe impairment in grammar than in vocabulary, as based on comparison of standardised scores in language tests, then pass through to the second stage.

In the second stage, children are administered a series of tests devised by van der Lely to assess the specific grammatical abilities that characterize G-SLI (van der Lely, 1996b, 1997c, 2000). Although standard tests assess a wide variety of skills within the area of syntax or vocabulary, van der Lely's tests target specific areas of grammar that children with G-SLI find particularly difficult – verb agreement and tense, reversible passives and pronominal reference. The three tests used are:-

- Verb Agreement and Tense Test (VATT; van der Lely, 2000) - designed to assess the expression of tense and agreement marking on a range of high and low frequency regular and irregular verbs. 20 past tense and 20 third person singular verbs are elicited.
- Test of Active and Passive Sentences (TAPS; van der Lely, 1996b) - a comprehension task designed to assess the assignment of theta roles in 12 active, 12 long passive, 12 short verbal passive and 12 short adjectival passive sentences.
- Advanced Syntactic Test of Pronominal Reference (ASTOP; van der Lely, 1997c) - a comprehension task designed to test reference of personal pronouns and anaphors. I used either the short (48 items) or long (96 items) form, depending on how much time was available during the testing session.

Note that although the tests have not yet been standardized, G-SLI participants make more than 20% errors on each of these tests, whereas normally-developing children rarely make errors after 6-8 years of age (van der Lely, 1996a; van der Lely & Stollwerck, 1997).

Details of G-SLI children's performance on these three tasks, plus their scores on the TROG, BPVS, non-verbal IQ, and their age at the first experimental testing are presented in Table 2.1. The TROG tests receptive grammar, with constructions such as reversible passives, comparatives, singular versus plural nouns, relative clauses and embedded sentences. The BPVS tests receptive vocabulary knowledge of a range of nouns, verbs and adjectives. The raw scores from the TROG and BPVS are used for matching to control children (see Section 2.3). Various IQ measures are used, reflecting work that has been going on over the past few years at the Centre for Developmental Language Disorders and Cognitive Neuroscience to find a non-verbal IQ test that really does not rely on language. Some children have been tested on the RPM (Raven, 1998), some on the Block Design and Matrices subtests of the BAS (Elliott, 1996), and some on the RPM and the Block Design subtest of the BAS. For confidentiality reasons, children are identified by a two-letter code rather than by their name, and when I discuss individual children's performance in the text, these are the codes that I use[3].

[3] Note that for some children not all of van der Lely's three tests have been administered. These are children who were recruited for Froud and van der Lely's (in prep.) word-learning study in 2000-2001. They were selected to the subgroup on the basis of their performance on various standardised language measures, as well as the VATT. Some of these children participated only in the past tense judgement study in Chapter 6. Due to restrictions on testing time imposed by their school, and the fact that several have since left the group, there was no opportunity to test them on the TAPS and A-STOP.

Table 2.1. Details of G-SLI participants' age at first testing, and scores on language and non-verbal tests

Code	Age	TROG (z-score)	BPVS (z-score)	NVIQ (z-score)	VATT (% correct)	TAPS (% correct)	A-STOP (% correct)
BD	8;11	-2.27	-2.10	-0.07^{R+B}	12.50	39.58	54.17
DT	9;07	-1.15	-0.40	0.20^{R}	35.00	58.33	n/a
SA	9;07	-1.73	-1.53	-0.53^{R}	67.50	79.17	72.92*
QC	9;11	-1.73	-1.93	-0.93^{R}	10.00	64.58	43.75*
HD	10;05	-2.40	-1.53	-0.90^{BM}	n/a	60.42	n/a
CT	11;02	-2.07	-2.60	-0.40^{R}	2.50	62.50	47.92*
PR	11;06	-1.53	-1.53	-0.16^{R+B}	35.00	n/a	n/a
GS	11;08	-2.27	0.80	-0.85^{BM}	10.00	72.22	58.33*
SL	11;10	-2.47	-1.90	0.03^{R}	42.50	30.56	70.83*
LJ	12;00	-1.20	-1.20	0.53^{R+B}	72.50	54.17	80.56
OD	12;02	-2.07	-2.50	-0.23^{R}	12.50	36.11	68.75*
MS	12;08	-1.20	-1.60	-0.87^{R}	42.50	n/a	n/a
WS	13;00	-1.20	-0.13	-0.20^{R}	80.00	n/a	n/a
DD	13;03	0.00	-1.47	0.03^{R+B}	72.50	n/a	n/a
SM	13;04	-1.30	-1.60	1.15^{R+B}	12.50	n/a	n/a
TD	13;04	-2.07	-1.87	-0.77^{R+B}	52.50	n/a	n/a
RP	13;06	-1.20	-1.33	-1.00^{R}	20.00	58.33	79.17
GD	13;11	-0.70	-2.20	-0.71^{R+B}	27.50	62.50	n/a
BS	14;11	-2.73	-2.47	-1.00^{R}	2.50	47.92	n/a
KA	15;05	-1.60	-2.80	0.18^{R+B}	67.50	n/a	n/a
LM	16;00	-1.60	-2.00	-0.40^{R+B}	15.00	n/a	n/a
DA	16;01	-0.73	-1.70	-0.27^{R+B}	37.50	n/a	n/a
CM	16;03	-0.70	-2.00	-0.07^{R+B}	60.00	n/a	n/a
PC	16;07	-1.20	-2.10	0.13^{R+B}	60.00	n/a	n/a
Mean (SD)	**12;09** (2;03)	**-1.55** (0.67)	**-1.65** (0.81)	**-0.30** (0.54)	**36.96** (25.32)	**55.67** (14.66)	**62.15** (13.06)
Range	107–199	-2.73–0.00	-2.80–0.80	-1.00–1.15	2.50–80.00	30.56–79.17	43.75–80.56

^{R} = Raven's Progressive Matrices (Raven, 1998); ^{BM} = Block Design and Matrices subtests from the British Ability Scales (Elliott, 1996) $^{R+B}$ = Composite score of Ravens and the Block Design subtest from the British Ability Scales

* = Short A-STOP

G-SLI children are characterized as having a primary impairment in grammar, which has secondary effects on the lexicon. One might therefore expect to see that these children's standard scores on tests of grammar would be lower than their standard scores on tests that measure lexical skills. This is true for some participants, but not for all, and not for the group as a whole. I discuss some reasons why this might be so:-

- Potential G-SLI children are tested on the TROG-1, which is only standardised up to the age of 12;11, and some of our G-SLI participants are older than this (note that the TROG-2, which is standardised up to the age of 16;11, was not available at the time of testing). In addition, the TROG has little discriminating ability beyond the age of about 9;00 or in the higher range of raw scores 16-20, where ceiling effects are seen. I have used the standard scores for 12 year olds when calculating the standard scores for these individuals, with the understanding that this will probably give a higher score than reflects their true ability. Since G-SLI participants are matched to typically developing controls on *raw* scores this is not a problem. The BPVS is standardized up to 15;08, and so accurate scores can be obtained for all but the very oldest children.
- Only some of the constructions tested in the TROG are relevant to the RDDR hypothesis and the impairment in G-SLI (Bishop, Bright, James, Bishop & van der Lely 2000) – blocks H (reversible active), L (reversible passive), N (subject post-modified by a verb phrase or prepositional phrase), R (object modified by relative clause) and T (centre-embedded sentence). So it is not unexpected for a G-SLI child to achieve a raw score of 15/20 – close to the point where standardisation is unreliable.
- Our G-SLI participants receive intensive speech and language therapy at the special language schools and units they attend. Much of this therapy is targeted towards the specific constructions in the TROG. Furthermore, each construction occurs only four times in the TROG. The more specific diagnostic tests that van der Lely has designed for the identification of G-SLI test the child over a range of constructions that G-SLI children find particularly difficult, and contain a much higher number of tokens for each construction. It is therefore much easier for G-SLI subjects to get a high score on the TROG than it is to achieve a high score on van der Lely's tests.
- The BPVS tests lexical recognition. Many different linguistic skills contribute to vocabulary knowledge, including semantics (Bloom, 2000). Chomsky considers semantics to be part of the computational component of language (Chomsky, 1993). The semantic abilities of G-SLI children have not yet been explored, and it may be that a deficit in this area impacts on BPVS scores. Vocabulary acquisition also relies on syntactic bootstrapping (Bloom, 2000), and syntax is impaired in G-SLI. Therefore,

despite the proposed dichotomy between grammar and the lexicon (e.g. Pinker, 1999), there is no reason to presume that children's lexical abilities are not impacted by their grammatical difficulties (van der Lely, 1999).

Note that the studies performed for this thesis took place at three different testing sessions, over a period of 22 months. Although it would have been desirable to have exactly the same group of children at all three sessions, this was not possible in practice for various reasons – children drop out of long-term projects, and others are recruited to the group once the first or second stage of testing is complete; sometimes a school will not allow a researcher to visit because they are worried about children being over-tested; sometimes children are ill and the visit cannot be re-arranged; and occasionally experimenter error results in a particular child's data not being recorded. All these factors conspire against obtaining a full set of experimental results for every child.

In Table 2.2 I list the children who undertook each task. In each chapter, in the section devoted to participant details, I detail the characteristics of just the group of G-SLI children who took part in that particular study, based on the most up-to-date scores of language-tests used for matching purposes, and so which may be different to the ones given in Table 2.1.

Table 2.2. Participants in each task

Task	Chapter	G-SLI children	Testing phase
Past tense elicitation	3	Data analysed from a different set of children (van der Lely & Ullman, 2001)	n/a
Past tense elicitation	4	BD, CM, CT, GD, GS, HD, LJ, OD, QC, RP, SA, SL, SM, WS	2
Non-word repetition	5	BD, CM, CT, GD, GS, LJ, OD, QC, SA, SM	1
Past tense judgement	6	BD, CM, CT, DA, DD, GD, KA, LJ, MS, PC, PR, LM, QC, SA, SM, TD	1
Past tense elicitation	7	BD, CM, CT, GD, GS, HD, LJ, OD, QC, RP, SA, SL, SM, WS	2
Past tense elicitation	8	BD, BS, DA, DT, GS, HD, KA, LJ, OD, QC, SL, SM, TD	3
Plural and present progressive elicitation	9	BD, BS, CT, DA, DT, GS, HD, KA, LJ, OD, QC, SA, SL, SM, TD	3
Comparative/ superlative elicitation	10	BD, CM, CT, GD, GS, HD, LJ, OD, QC, SA, SL, SM	2
Adjective-from-noun elicitation	11	BD, CM, CT, GD, GS, HD, LJ, OD, QC, SA, SL, SM	2

2.3. Selection of control groups

The G-SLI group's performance is compared to that of typically developing children. The reasons for this are two-fold: first, to get picture of normal development and second, to match individual control children with G-SLI participants. The best choice of suitable control matches is not clear cut. There is little to be gained from using chronological age matches for language tasks because SLI children will by definition perform more poorly (although age-matched controls are appropriate for non-linguistic tasks). Using language-matched controls is more informative, because it allows us to determine whether poor performance on the experimental task is to be expected given the G-SLI children's general low language level, or whether the experimental task has identified an area of deficit above and beyond that expected for their language level (the 'delay within a delay' model of Rice, 2003, alternatively termed the 'delay and disruption' pattern, Rice, 2004).

The value in discovering which aspects of language are delayed above and beyond what is expected is two-fold. It enables us to (1) to explore 'little modularity' (see Section 1.2.1), i.e. to identify which components of the language system can be differentially impaired, and (2) to discover what possible causes of SLI are, with the assumption that the area(s) of language that is the most impaired constitutes the core deficit(s).

Control participants were all chosen to have English as a first language, no history of a speech and language disorder, and no history of a hearing impairment. Two tests were used for matching – the TROG, to provide grammar matches, and the BPVS, to provide vocabulary matches. In previous work, van der Lely's control groups have comprised typically developing children aged 5;05-8;09, divided into two or three age bands (e.g. van der Lely & Stollwerck, 1997; van der Lely & Ullman, 2001). The youngest of these groups are matched for grammar abilities, while the second (and third, if used) group are matched on vocabulary abilities. This allows a match for different aspects of language ability. It also enables the developmental picture for typically developing children to be investigated. It should be stressed that these are group matches rather than individual matches.

One of the experiments that I carried out in the first testing phase for this thesis (presented in Chapter 6) uses *individual* matches. There are two reasons for choosing individual matches. First, the age range and language range of the G-SLI children is much wider than that of the G-SLI groups used in van der Lely's previous studies: therefore it is arguably not appropriate to use a narrow age-group of typically developing matches when assessing statistical differences between the groups on tasks with quantitative measures. Secondly, it is not clear whether all van der Lely's control groups really are matched for the aspect of language that they are claimed to be matched on. For example, in her study of passives, van der Lely claims that the G-SLI group are not significantly different from the two older control groups (termed LA6 and LA7, aged 6;05-8;09 years) on naming vocabulary, when in fact the p value is only 0.087 (van der Lely, 1996a). There are concerns over such low p values when two groups are being matched. Mervis and Robinson (2003) write that 'Rejecting the null hypothesis because it is improbable under the theoretical sampling distribution should be a very different decision from accepting the null hypothesis because it is likely to be true.' (p. 236). They recommend that any value less than $p = 0.20$ (2-tailed) is too low to accept the null hypothesis that there is no significant difference between the two groups, and any value between 0.20 and 0.50 is ambiguous. P values of 0.50 and over should be obtained before the groups are considered as matched. Individual matching means that if controls really are matched on identical raw score, then p value will be close to 1.00, and the group means, ranges and standard deviations will be identical. Such matching criteria provide a better basis for comparisons. More seriously, in the control groups used in van der Lely and Christian's (2000) study, neither of the two control groups is matched for vocabulary, even though compounding is a word-building task.

However, the issue still remains as to how G-SLI participants' performance compares to the normal developmental profile. It is valid to ask whether the group as a whole, or individuals in the group, are performing at a similar level to 6, 7 or 8 year old typically developing children. An even more important consideration is whether the pattern of linguistic behaviour is the same as that of younger typically developing children. These questions can only be answered by looking at groups of typically developing children divided into age bands one or one and a half years apart. I therefore use this method of matching in experiments from the third phase (reported in Chapters 8 and 9). I use a hybrid method of matching in one of the experiments from the first phase of testing (Chapter 5) and in the experiments of the second testing phase (reported in Chapters 4, 7, 10 and 11), whereby I select individual matches and then divide the children into two groups according to age.

One inherent disadvantage in using language-matched controls is that there is a large chronological age discrepancy between them and the G-SLI participants. Not only language is needed for the experimental tasks, but also more general cognitive skills such as memory and attention. This raises the possibility that G-SLI and their language matches might perform at similar levels on a particular task for different reasons – the G-SLI children because of their poor language skills, and their controls for their immature cognitive abilities. Using language matches is therefore a conservative method where it comes to finding significant differences in group means (Bishop, 1997:239).

2.4. Controversy over G-SLI

Joanisse (2004) and Tomblin and Pandich (1999) argue that children with G-SLI are at the lowest end of a normal distribution of grammatical abilities. Therefore, contrary to van der Lely, they claim that G-SLI does not exist as a phenotypically and genotypically separate subgroup. This claim appears to be based, at least in part, on misinterpretations of van der Lely's data and the conclusions she and her colleagues draw from them. For example, Joanisse stresses that a rule-deficit account (e.g. Pinker, 1991) would not predict van der Lely and Ullman's (2001) finding that irregular past tense forms are impaired to the same level as regular past tense forms, and he says that a delay in irregulars is also found in typically developing children. However, what van der Lely and Ullman's data show (and I replicate this finding in Chapter 8) is that typically developing children show an advantage for regular verbs over irregulars, whereas G-SLI children show no such regularity advantage. Therefore van der Lely and Ullman are justified in interpreting their data as showing that regulars are affected more than irregulars. Tomblin and Pandich (1999) find no children who have grammatical deficits but no vocabulary deficits in their studies, but in fact van der Lely has never claimed that G-SLI children show a profile of poor grammar but intact vocabulary. Instead she contends that lexical learning requires, among other skills, the ability to use syntactic cues, and that difficulties in lexical learning could be secondary to the syntactic deficit (van der Lely, 1999).

Regardless of the validity of G-SLI as a subgroup, the point remains that van der Lely and colleagues have strict criteria for their choice of participants, and this detailed characterization remains essential. The issue of whether G-SLI children are at the tail-end of a normal distribution of grammar abilities, or a subgroup with grammatical behaviour different to the norm, is one that does not impact on the value of the linguistic findings presented in this thesis. The work in this thesis provides a detailed phonological phenotype of G-SLI, particularly as regards the impact of phonology on morphology. Establishing precise phenotype/genotype relations is an essential pre-requisite for studies seeking to identify the genetic basis of SLI (e.g. Lai, Fisher, Hurst, Vargha-Khadem & Monaco, 2001; SLI consortium, 2002). Ultimately, it may only be through understanding the genetic and neurological underpinnings of SLI that we can settle the issue of whether G-SLI children are qualitatively different from children with normally-developing language and from other SLI children. Comparing the genetics of different subgroups may provide an explanation for the heterogeneity in SLI, as it is possible that the observed phenotypic variability is a result of genetic variation (van der Lely, 1999; van der Lely & Stollwerck, 1996).

A detailed investigation of language is also valuable because it may be that when language structures break down, they break down in similar ways whatever the deficit, be it SLI, language delay, Down's Syndrome, Williams Syndrome etc. The linguistic findings from the G-SLI population, and the tools developed for probing language structures, may then help advance our knowledge of language deficits in other developmental disorders. Relevant to this issue is the existence of a conflict between clinical and theoretical aims when selecting SLI children for research. Clinically-orientated research can use inclusive criteria in participant selection because impaired language may require the same therapy whether or not children have co-occurring deficits such as low non-verbal IQ, Asperger's Syndrome, dyspraxia, or Down's Syndrome. For theoretically-oriented research, however, the inclusion of children with co-occurring deficits makes it harder to distinguish causal factors. For example, if children with both low non-verbal IQ and SLI are studied, we cannot rule out the possibility that the group's poor performance on a language task is the result of low IQ.

In addition to the clinical/theoretical divide is a further conflict of aims: those of the cognitive sciences versus those of linguistics. Whereas developmental cognitive science is interested in what population as a whole does, current linguistics is inspired by the Chomskyan notion of an individual grammar, whereby studying what goes on inside the head of any particular speaker is of value in exploring the boundaries of grammatical knowledge. Therefore a tension exists when researching language disorders between using pure groups such as G-SLI and using more inclusive groups: the more inclusive the group, the larger the number of children that can be studied, but such research tends to miss out on the linguistic details. At the opposite extreme, concentrating on the detailed linguistic behaviour of one or two children runs the risk that this behaviour is rare and not representative of the wider population. The work in the chapters that follow tries to maintain a balance between investigating the characteristics of the G-SLI group as a whole, while also commenting on the individual linguistic behaviour of individuals, and how this can inform theories of linguistics and cognitive science. I do not pretend to have got the balance right, but I believe that the framework I set out has benefits for both levels of analysis.

PART 2.

Establishing the morphological and phonological deficits

Chapter 3. Establishing the morphological deficit: The impact of verb-end phonotactics on regular past tense inflection

3.1. Introduction

3.1.1. Chapter outline

In this chapter I show that G-SLI children have a morphological deficit that is independent of phonology. I do this by investigating the impact of verb-end cluster phonotactics, in a reanalysis of elicitation data from van der Lely and Ullman (2001).

I begin in Section 3.1.2 by summarising van der Lely and Ullman's (2001) study of past-tense morphology in G-SLI, and discuss some of the criticisms levelled at their interpretation of the results. In Section 3.1.3 I consider the phonotactics of regular inflected verbs and set out a typology of what I term 'legal' and 'illegal' clusters. In Sections 3.1.4, 3.2 and 3.3 I reanalyse van der Lely and Ullman's data with regards to cluster phonotactics, and I show that the G-SLI group have greater difficulty inflecting verbs when an illegal cluster would be formed. In Section 3.4, I make predictions regarding the impact of phonotactics on past tense inflection in individuals with a different developmental disorder – Williams Syndrome (WS) – and test these predictions using data from Thomas *et al.*'s (2001) study. I argue that individuals with WS are able to use phonotactics as a cue to morphological complexity. I show in Section 3.5 that van der Lely and Ullman's original interpretation of their data, that there is a morphological deficit in G-SLI, is correct. Importantly, these phonotactic data are indicative of a morphological deficit rather than a phonological one.

3.1.2. Past tense morphology in G-SLI

Van der Lely and Ullman (2001) investigated regular and irregular past tense formation in a group of twelve children with G-SLI (age range 9;03-12;10) and in three groups of typically developing children matched on standardised measures of morphology and vocabulary. They found that the G-SLI and language ability (LA) controls showed quantitatively and qualitatively different patterns of performance. The G-SLI children's production of regular past tense forms was significantly lower than that of all three control groups. Their performance on irregular past tense verbs was lower than that of the two groups matched for vocabulary level (LA2 and LA3) but not significantly different from the group matched for morphology (LA1). These findings suggest that the difficulties G-SLI children face with regular inflection cannot be solely accounted for by their language age. Two patterns in the data are crucial to van der Lely and Ullman's interpretation of the data.

- The control children showed a significant advantage for regular over irregular verbs, whereas the G-SLI children didn't.
- The G-SLI children showed a consistent frequency effect for both regular and irregular verbs, whereas the controls do so only for irregular verbs.

Van der Lely and Ullman claim that a dual mechanism model of morphology (e.g. the 'Words and Rules' model, see Section 1.2.2) offers a parsimonious explanation for these results. They conclude that (1) both groups of children retrieve irregular past tense forms from the lexicon and (2) G-SLI children retrieve stored regular forms from the lexicon, whereas typically-developing children compose regular forms *de novo* from the verb stem and the *-ed* suffix. In other words, G-SLI children have a morphological deficit, which results in them being inconsistent in their use of the regular suffixation rule, and so they have to rely on the storage of forms that are already inflected.

The dual mechanism interpretation has been criticised by proponents of a single mechanism model, who claim that a phonological impairment underlies the difficulty with regular morphology. At the heart of the single mechanism explanation is that G-SLI children have a deficit in the processing of rapid auditory stimuli, which particularly impacts on the non-salient */t/* and */d/* inflection, resulting in poor phonological representations of past tense forms. Studies by van der Lely, Rosen and Adlard (in press) demonstrate that G-SLI children have no consistent auditory deficit, and that performance on tasks involving auditory discrimination does not correlate with phonological and other language abilities. However, the proposal that some aspect(s) of regular past tense phonology affect G-SLI performance is worth investigating, which is why I take it up here and in subsequent chapters.

3.1.3. The phonological characteristics of past tense forms

The morphological processes involved in regular and irregular inflection are traditionally analysed as being different (e.g. Kiparsky, 1982). The addition of the past tense suffix is morphology at the word level, while irregular past tense forms are instead created by root-level morphology. There is an important phonological difference between these two levels – while root-level morphology creates forms with word-endings that are phonotactically identical to monomorphemic (i.e. uninflected) English words, word-level morphology *can* give rise to word-endings which are phonotactically different to those in monomorphemic words. The phonotactic sequences found in regularly inflected words arise by lexical insertion, by 'accident' from a phonological viewpoint as it were (Harris, 1994). Irregular past tense forms tend not to contain clusters, e.g. *took*, *swam*, *got*, *stood*, but if they do, e.g. *slept*, *built*, *spent*, *lost*, then those clusters are legal in English, meaning that they also occur in monomorphemic words (c.f. *accept*, *stilt*, *tent*, *frost*). Some regularly inflected forms also contain legal clusters, e.g. *crossed* (c.f. *frost*, *mist*), *scowled* (c.f. *bald*, *cold*), *dropped* (c.f. *opt*, *apt*). However, many regular past tense forms have clusters which can *not* occur in monomorphemic words, and are therefore illegal in English phonology e.g. *slammed*, *rushed* and *changed*. This distinction between phonotactically legal and illegal sound sequences is justified by Harris' convincing discussion that one can only determine what is phonologically possible in a language by confining one's data to morphologically simple words (Harris, 1994).

Table 3.1 presents a comprehensive typology of irregular and regular inflected verb endings in English (assuming a non-rhotic accent). The table distinguishes between short and long vowel length for reasons that will become clear in the discussion of nucleus + rhymal consonant phonotactics that follows. It also distinguishes between the voiced and unvoiced suffix. The syllabic suffix, which is added to stems ending in /*t*/ and /*d*/, is not considered here because it does not create verb-end clusters. V stands for a short (lax) vowel, VV for a long (tense) vowel, and O for an obstruent and S for a sonorant. Monomorphemic forms are presented, where possible, beneath each verb. A '_' indicates that a particular past tense form is not attested, and a '*' indicates that a particular monomorphemic form does not occur. In bold are those regular verbs that do not have a monomorphemic counterpart, and are hence 'illegal'.

Table 3.1. Irregular and regular inflected verb endings in English

	Irregular		Regular	
a	VV*d*	*made, rode*	VV-*d*	*allowed, spied, rowed*
		maid, road		*loud, wide, code*
b	VO*t*	*kept, left, lost*	VO-*t*	*capped, packed, sniffed*
		adept, left (adjective), *frost*		*apt, pact, lift*
c	VO*d*	_, _, _	VO-*d*	***robbed, hugged, judged***
		*, *, *		*, *, *
d	VS*t*	*felt, spent*	VS-*t*	_, _
		felt (noun), *tent*		*felt, tent*
e	VS*d*	*held*	VS-*d*	*yelled, conned*
		weld		*weld, pond*
f	VVO*t*	_, _, _	VVO-*t*	*paced*, ***seeped, peeked***
		paste, *, *		*paste*, *, *
g	VVO*d*	_, _, _	VVO-*d*	***caged, dived, raised***
		*, *, *		*, *, *
h	VVS*t*	_	VVS-*t*	_
		paint		*paint*
i	VVS*d*	*told, found*	VVS-*d*	*rolled, drowned*
		old, round		*cold, sound*
j	VOO*t*	_	VOO-*t*	*flexed*
		next		*next*
k	VOO*d*	_, _	VOO-*d*	_, _
		*, *		*, *
l	VSO*t*	_	VSO-*t*	*stomped*, ***winced, helped***
		prompt		*prompt*, *, *
m	VSO*d*	_, _,	VSO-*d*	***plunged, bronzed***
		*, *		*, *
n	VVSO*t*	_, _	VVSO-*t*	***pounced***
		*, *		*
o	VVSO*d*	_,	VVSO-*d*	***changed***
		*,		*

Let's first consider the phonotactics of monomorphemic words that end in a coronal stop (i.e. */t/* or */d/*). There are several points of note:-

1. Clusters of voiced obstruent + */d/* are illegal, whatever the length of the preceding nucleus (see c, g).
2. Clusters of unvoiced obstruent + */t/* are illegal when the preceding nucleus is long, with the exception of */s/*, which shares its place of articulation with */t/* (see f).
3. Final */t/* can support larger clusters than final */d/*, although words with 3-consonant clusters are rare. 3-consonant clusters are confined to */mpt/*, e.g. *tempt*, *prompt*, */kst/*, e.g. *next*, *text*, and */lkt/*, only example *mulct*.

Now let's consider the phonotactics of irregular verbs. If a particular verb-end pattern is not legal in a monomorphemic form, then that pattern will not appear in an irregular verb. There are certain gaps where patterns that we would predict to occur in irregulars, e.g. VVS*t* (see h), are not attested. These lexical gaps are surely due to historical accident, as there is no principled phonological reason why they should not occur. With only 150-180 irregular verbs (Pinker, 1999), it is not unexpected that some predicted forms are not found.

In terms of the phonotactics of regular past tense forms, there is nothing to stop a suffix being attached to any regular verb stem, regardless of its stem-end phonology. Any gaps that occur in Table 3.1 are due to suffix having to agree in voicing with the stem-final consonant. Therefore, */t/* is not permissible after a sonorant (see d and h). */d/* is not permissible after an obstruent/obstruent cluster because such clusters must be unvoiced in English (e.g. */ask/*, **/azg/*), and therefore the suffix must be */t/* (see k). The constraint against having a voiced obstruent cluster derives from the properties of speech - it is more difficult to maintain vocal cord vibration when there is a constriction of the type that produces a fricative or an oral stop. This constraint, formulated in various ways, e.g. NOTTWICE(+VOICED) (Bernhardt & Stemberger, 1998), is active in many languages, such as English, but dominated in other languages, such as Polish, which allows such sequences to occur monomorphemically.

One might ask why the illegal outputs of regular inflection are tolerated at all. If a voiced obstruent cluster such as */gd/* is not legal morpheme internally, why should it be permissible in a past tense form? The answer must surely lie in paradigm identity i.e. *tugged* is related to *tug*, and if it changed to */tʌkt/*, it would lose identity with *tug* and be indistinguishable from the past tense of *tuck*. It seems plausible that the phonological well-formedness constraints, which apply to monomorphemic forms, do not in this instance apply to the output of word-level morphology because the meaning of the output would be less transparent (Paradigm identity is a well-discussed phenomenon: see McCarthy, 2004, for an up-to-date discussion).

It is important to note that illegal phonotactic sequences occur *only* when morphology has taken place. This is relevant when considering the terms 'legal' and 'illegal'. Joanisse (personal communication, March 2003) cautions against using 'illegal' because that term suggests that a particular sequence cannot occur in English when it patently does. He cites *borscht* and *damned* as examples of words with illegal clusters – yet the former is loaned from Polish (and has an alternative pronunciation where the cluster is simplified by omission of the /t/), while the latter is morphologically derived. Foreign words are irrelevant to the phonology of English – indeed, loan words into any language are eventually assimilated into the phonological patterns of that language (indeed, *borscht* is already simplified from *borshch*, whose final /ʃtʃ/ cluster does not even occur in derived words of English). The crucial point is that loan words aside, illegal clusters occur only where there is morphology. Joanisse suggests using 'marked' and 'unmarked' instead, but in my view that reveals a fundamental misunderstanding of the phenomenon. It is misleading to say that voiced obstruent clusters are marked in English because this implies that they do exist word-finally, they are just disfavoured in some way. Languages vary in how they rank constraints, and in English NOTTWICE(+VOICED) is ranked high enough to be active. It is not that voiced obstruent clusters are disfavoured in some unprincipled way – they can *only* occur when morphology has taken place. They can occur internally in words which have originated through compounding (e.g. *humdrum*, *ashtray*) or prefixation (e.g. *abduct*). They can also occur across word boundaries (e.g. *Tom did*). When they occur word-finally, they unambiguously signal past tense inflection.

Stemberger prefers the terms 'basic' and 'non-basic' for legal and illegal clusters respectively (personal communication, March 2003). These terms are preferable to 'unmarked' and 'marked', but I continue to use 'legal' and' illegal', as these communicate much more strongly that certain sequences do not occur monomorphemically.

3.1.4. Establishing the status of morphology in G-SLI using verb-end phonotactics

I propose that investigating the impact of phonotactics on regular past tense formation in both G-SLI and typically developing children provides a unique test of the single mechanism and WR accounts, as it enables us to determine whether a morphological rule is required for regular forms. Although the predictions of a single mechanism account regarding the impact of phonotactics on regular inflection have not previously been articulated, a coherent prediction is for regular verbs with legal clusters to be acquired earlier and produced more easily than those with illegal clusters because of their relative frequencies: legal clusters are more frequent because they occur in both monomorphemic and inflected forms, whereas illegal clusters are less frequent because they only occur in inflected forms. So both typically developing children and those with G-SLI are expected to perform better on regular verbs containing legal clusters than on those containing illegal clusters.

As for the predictions of the WR model, the role of phonotactics has likewise not been addressed. I propose that the WR model differs from the single mechanism account in the prediction it makes for typically developing children. There are in fact two logical predictions:-

- Traditionally the WR model has claimed that phonology has no impact on rule use. Under this view, cluster phonotactics and the difference in frequency between legal and illegal clusters should not affect performance.
- A further prediction recognises that illegal clusters signal morphological complexity. The illegal clusters of *slammed*, *robbed* and *rushed* indicate that these words can only be past tense verbs, whereas words with legal clusters are ambiguous between inflected and uninflected forms (e.g. */mɪst/* is both the past tense of *miss* and the noun *mist*). It is possible that paradigms that introduce illegal clusters are more easily learnt by children who are able to use morphological rules than by those who aren't. Therefore children who are old enough to have acquired all the clusters involved might experience greater success on regular verbs with illegal clusters.

For G-SLI children, the WR model predicts that if they have an impaired morphological rule and are relying on the phonological storage of regular forms, then cluster frequency will play a crucial role. Therefore G-SLI children are expected to perform better on legal regulars because of their higher frequency, just as the single mechanism account predicts.

It should be stressed that little is currently known about the sequence of word-final cluster acquisition (Bernhardt & Stemberger, 1998; Kirk & Demuth, 2003). It seems likely that cluster frequency will play a role, with more frequent clusters being acquired before less frequent, although this is of course a generalisation that abstracts away from the segmental characteristics of individual clusters (Stemberger, personal communication, March 2003).

3.2. Method

3.2.1. Procedure

The data are taken from van der Lely and Ullman (2001). Of the ten regular verbs selected for this analysis, five have an inflected form with a legal cluster and five have a cluster that is illegal by the criteria discussed in Section 3.1.3. The two groups of verbs are matched for cluster complexity (two consonants) and mean past tense frequency (see Table 3.2). These frequency counts were drawn from the British English COBUILD corpus of the University of Birmingham, by the Centre for Lexical Information (CELEX) at the University of Nijmegen. Individual verb frequencies were augmented by 1 and ln-transformed. Care was taken to match the two groups for frequency because van der Lely and Ullman showed that for G-SLI children, but not for their language-matched controls, there is a significant correlation between past tense frequency and correct past tense inflection.

In addition, I used the CELEX database to calculate the frequencies of verb-end clusters. Each cluster frequency was calculated by summing the past tense frequencies of all the verbs that contain that cluster, and for legal verbs all word-final monomorphemic instances of that cluster were also counted. The verbs used are shown in Table 3.2, along with their past tense frequencies and cluster frequencies. It is evident from Table 3.2 that the legal clusters are much more frequent than the illegal clusters, with no overlap in distribution.

Table 3.2. Verbs used in this analysis

Legal verb-end clusters			**Illegal verb-end clusters**		
Verb	**Past tense frequency**	**Cluster frequency**	**Verb**	**Past tense frequency**	**Cluster frequency**
scowled	2.3	8.14	*tugged*	2.9	3.69
flapped	2.6	6.53	*slammed*	3.6	5.76
stalked	2.7	7.82	*rushed*	4.4	5.23
crossed	5.1	9.24	*robbed*	3.1	4.16
dropped	5.6	6.53	*flushed*	3.9	5.23
Mean (SD)	**3.66** (1.56)	**7.65** (1.15)	**Mean** (SD)	**3.58** (0.61)	**4.81** (0.85)

3.2.2. Predictions

To recap, the single mechanism account predicts that typically developing children and those with G-SLI will inflect verbs containing legal clusters more successfully than verbs with illegal clusters. The WR model makes this prediction for children with G-SLI, but makes different predictions for typically developing children, namely that for them performance on both types of verbs will be the same, or alternatively that those with illegal clusters will be easier.

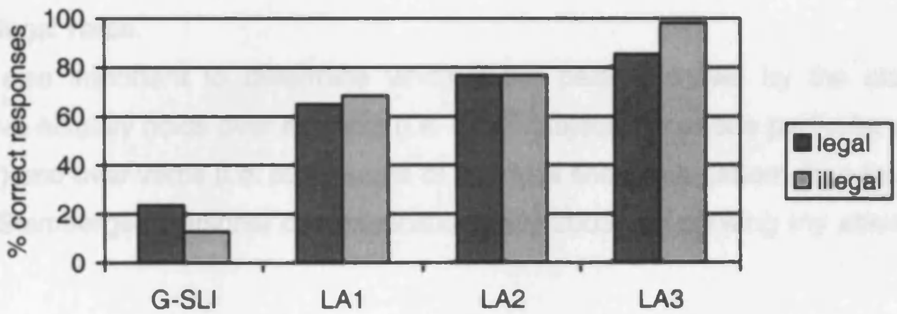
3.3. Results

The data presented in Table 3.3 and Figure 3.1 show that the G-SLI group performs worse than all three control groups on both legal and illegal verbs. Because the data are not normally distributed and did not respond to transformation, I used non-parametric tests in the analysis.

Table 3.3. % correct responses for verbs with legal and illegal verb-ends

Condition		G-SLI 9;03-12;10	LA1 5;05-6;04	LA2 6;05-7;04	LA3 7;05-8;09
Legal	**mean (SD)**	23.63 (17.48)	65.00 (30.90)	76.67 (28.07)	85.40 (25.01)
Illegal	**mean (SD)**	12.73 (13.48)	68.33 (30.10)	76.67 (25.35)	98.18 (6.03)

Figure 3.1. % correct response for verbs with legal and illegal verb-ends



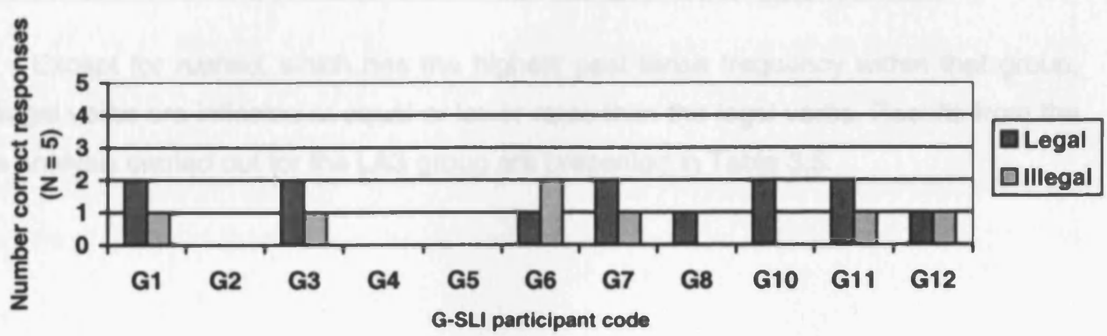
First I investigated the developmental profile across the three groups of control children. A Wilcoxon Signed Ranks test on the control group as a whole revealed that illegal verbs were inflected more often than the legal verbs, although this did not reach significance, $Z = -1.461$, $p = 0.144$. However, a developmental difference was detected across the three groups. For the LA1 and LA2 groups legality had no significant effect on performance, $Z = -0.587$, $p = 0.557$, and $Z = -0.214$, $p = 0.831$ for LA1 and LA2 respectively. For the LA3 controls however, illegal verbs were significantly easier than the legal verbs, $Z = -2.121$ $p = 0.034$. As can be seen from Figure 3.1, these results show that for typically developing children aged 5;05-7;04, phonotactics do not significantly influence performance, whereas for children aged 7;05-8;09, illegal verbs are *easier* to inflect.

Next I consider the performance of the G-SLI group relative to that of typically developing children, and for this I collapsed the data from the three control groups. Kruskal-Wallis tests conducted on legal and illegal verbs separately revealed a significant effect of group for both verb types: for legal verbs, $\chi^2(1) = 17.751$, $p < 0.001$, and for illegal verbs, $\chi^2(1) = 22.107$, $p < 0.001$. These results show that the regular deficit that van der Lely and Ullman found in the G-SLI group relative to the language-matched control group holds even when the verbs are separated along the legal/ illegal dimension.

Finally, I investigated whether G-SLI and control children's responses were similarly affected by legal and illegal phonotactics. A Kruskal-Wallis by subject analysis was used to investigate the group (G-SLI, control) x difference between the illegal and legal scores (illegal score minus legal score). The interaction was significant, $\chi^2(1) = 5.461$, $p = 0.019$. Thus the G-SLI and control groups respond differently to phonotactic legality. A Wilcoxon signed ranks test comparing G-SLI performance on legal and illegal verbs revealed that the G-SLI group inflected illegal verbs significantly less often than legal verbs, $Z = -1.897$, $p = 0.029$ (1-tailed). Remember that, as reported above, this is the opposite pattern to the LA3 group, who inflected illegal verbs more successfully. It is also different to the performance of the LA1 and LA2 groups, who performed equally well on illegal as on legal verbs.

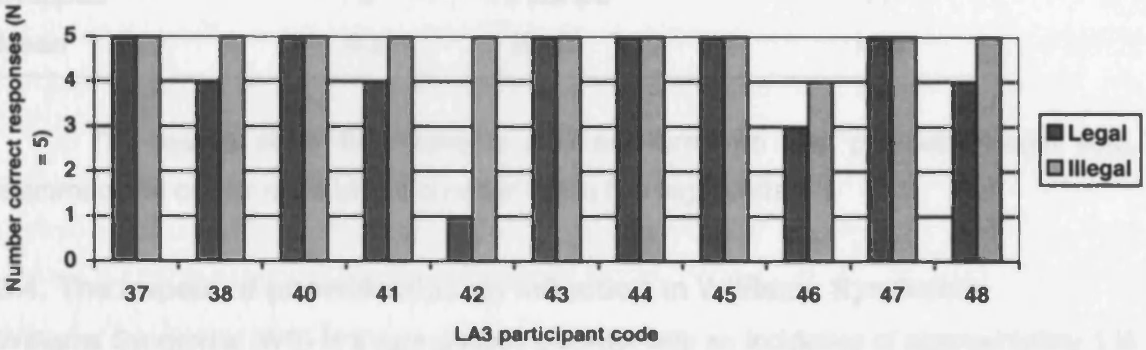
It is also important to determine whether the pattern shown by the statistical analysis above actually holds over subjects (i.e. more subjects show one particular pattern than another) and over verbs (i.e. more verbs of one type show one pattern than the other; I thank Joe Stemberger, personal communication, July 2003, for drawing my attention to this).

Figure 3.2. Performance by individual G-SLI children on legal and illegal verbs



For six of the G-SLI children, legal verbs are inflected more successfully than illegal verbs. For one child (G12) performance is equal, and just one (G6) finds the illegal verbs easier.

Figure 3.3. Performance by individual LA3 children on legal and illegal verbs



The same analysis performed on the LA3 group shows that for all 5 children whose performance is not at ceiling, verbs with illegal clusters are easier than those with legal clusters.

The number of G-SLI children successfully inflecting each verb is presented in Table 3.4.

Table 3.4. Number of G-SLI children successfully inflecting each verb

Legal	**Number (N=11)**	**Illegal**	**Number (N=11)**
Scowled	1	*Tugged*	1
Flapped	3	*Slammed*	1
Stalked	1	*Rushed*	4
Crossed	5	*Robbed*	1
Dropped	3	*Flushed*	0
Mean	**2.6**	***Mean***	**1.2**

Except for *rushed*, which has the highest past tense frequency within that group, the illegal verbs are inflected at equal or lower rates than the legal verbs. Results from the same analysis carried out for the LA3 group are presented in Table 3.5.

Table 3.5. Number of LA3 children successfully inflecting each verb

Legal	Number (N=11)	Illegal	Number (N=11)
Scowled	10	*Tugged*	11
Flapped	8	*Slammed*	10
Stalked	10	*Rushed*	11
Crossed	9	*Robbed*	11
Dropped	9	*Flushed*	11
Mean	**9.2**	**Mean**	**10.8**

The results show that there is only one error on one particular illegal verb, *slammed*, but one or more errors on each of the five legal verbs.

3.4. The impact of phonotactics on inflection in Williams Syndrome

Williams Syndrome (WS) is a rare genetic disorder with an incidence of approximately 1 in 20,000. It is caused by a microdeletion on one copy of Chromosome 7, which affects several genes. The syndrome is characterised by a specific physical, cognitive and behavioural phenotype. Within cognitive skills, verbal abilities are superior to visuo-spatial abilities, although language performance falls below that found in chronological age-matched controls. However, despite the relative strength of language skills, linguistic development is uneven, and there is dispute about actual performance on different tasks, let alone how to interpret this performance (Clahsen & Temple, 2003; Mervis, Morris, Bertrand & Robinson, 1999; Thomas *et al.*, 2001).

Pinker (1991, 1994) claims that WS (with higher verbal than non-verbal IQ) and SLI (with lower verbal than non-verbal IQ) together provide evidence of a genetic double dissociation between language and cognition. In addition, the two groups show different behaviour on regular as opposed to irregular past tense formation. Past tense data from Clahsen and Almazan (1998), using van der Lely and Ullman's (2001) task, reveal that the performance of a group of individuals with WS on regular verbs is equivalent to that of mental age-matched controls. Their performance on irregular verbs is, however, much lower than that of the controls. This is, of course, the opposite pattern to that found in SLI. However, Clahsen and Almazan's data have been criticised (Thomas *et al.*, 2001) because only 4 Williams syndrome subjects participated in the study, and because the control subjects performed at much higher levels on the irregular verbs than was the case in van der Lely and Ullman's own study. When Thomas *et al.* (2001) replicated the task with a larger group of 18 participants with WS and three groups of control children, the WS group did not show a selective deficit for irregulars.

Clahsen and Almazan (1998) propose that a selective deficit affects representation of, and access to, word-specific knowledge in WS. Specifically, the claim as it relates to past tense morphology is that irregular verbs, involving as they do word-specific knowledge, are selectively impaired, whereas the rule-based inflection of regular verbs is in line with mental age. In contrast, Karmiloff-Smith and Thomas (in press) defend a single mechanism approach, whereby the language in WS evolves according to an atypical balance of phonological and semantic constraints. They do not state whether the differential balance is brought about by a relative strength in phonology or a relative weakness in semantics. However, if under a dual route analysis morphological rules are indeed intact in WS, then with regards to the inflection of regular verbs we would not expect any advantage for legal clusters over illegal. In fact, we can go further to predict that if WS individuals actually overapply the regular inflection rule, as Clahsen and Almazan (1998) claim, then they may well be particularly sensitive to illegal phonotactics, and may actually perform better on verbs which produce illegal clusters. It is this specific prediction of better performance on illegal verbs that I test in the analysis presented here.

Michael Thomas and Annette Karmiloff-Smith kindly made available to me the data that they and their colleagues collected on WS individuals using van der Lely and Ullman's (2001) task (Thomas *et al.*, 2001). I use this data to test the prediction that the WS group will show an advantage for illegal verbs. Thomas *et al.* tested 18 children and adults with WS, range 10;11-53;03, mean 22;08. They also tested three groups of typically developing children, the youngest of which provides the best control in terms of past tense performance. The 10 children in this group ranged from 5;05-6;04, mean age 6;00. The results are shown in Table 3.6.

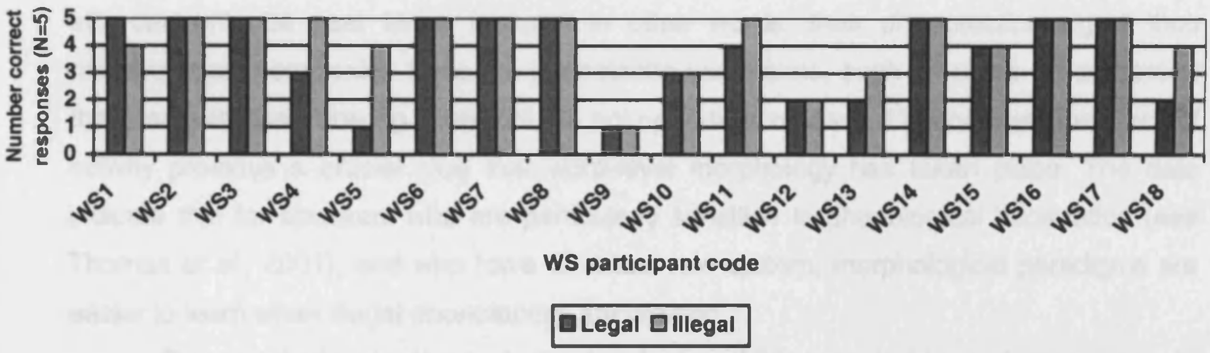
Table 3.6. % correct responses for verbs with legal and illegal verb-ends

Condition		Williams Syndrome	6 yr old controls
Legal	**Mean (SD)**	74.44 (30.53)	76.00 (35.02)
Illegal	**Mean (SD)**	83.33 (24.01)	80.00 (28.28)

I investigated whether WS and control children's responses are similarly affected by legal and illegal phonotactics. A Kruskal-Wallis test was used to check for an interaction between group (WS, control) and legality (as measured by the difference between the illegal and legal scores for each subject). The interaction was not significant, $\chi^2(1) = 0.370$, p = 0.543. Thus the WS and control groups respond similarly to phonotactic legality. A Wilcoxon signed ranks test comparing control performance on legal and illegal verbs revealed that although numerically performance is better on illegal verbs, this difference is not significant, Z = -0.816, p = 0.414. The same test carried out on the WS group, however, revealed that, as predicted, the WS group inflected illegal verbs significantly more often than legal verbs, Z = -1.807, p = 0.036 (1-tailed). Note that this is the opposite pattern to the G-SLI group, who inflect legal verbs more successfully.

An analysis was carried out by individual WS participant and then by verb. The results of this analysis are shown in Figure 3.6 and Table 3.7.

Figure 3.6. Performance by individual WS participants on legal and illegal verbs



For twelve participants performance is equal on legal and illegal verbs, and for eight of those performance is at ceiling in any case. Five participants show better performance on illegal compared to legal verbs (WS 4, 5, 11, 13, 18) and only one performs better on legal verbs (WS 1).

Table 3.7. Number of WS participants successfully inflecting each verb

Legal	Number (N=18)	Illegal	Number (N=18)
Scowled	12	*Tugged*	14
Flapped	13	*Slammed*	14
Stalked	12	*Rushed*	16
Crossed	15	*Robbed*	16
Dropped	15	*Flushed*	15
Mean	**13.4**	**Mean**	**15.0**

3.5. Discussion

3.5.1. Summary of the results

The three populations studied here – typically developing children, G-SLI children, and WS children and adults – show three qualitatively different patterns of performance on verbs with legal and illegal verb-end clusters. For typically developing children of the ages studied here (5;05-8;09), phonotactics either have no impact on inflection, or there is an advantage for verbs with illegal clusters. For G-SLI children, performance is significantly better on legal verbs, whereas for WS individuals, performance is better on illegal verbs.

3.5.2. The impact of phonotactics on past tense inflection

I propose that children and adults with WS find verbs with illegal clusters easier, despite their lower frequency, because their morphology is transparent. *Slammed*, *robbed*, *rushed* etc. can only be past tense forms – in other words, their phonotactics signal their morphological complexity. Because phonotactic constraints, such as place agreement of the nasal with the following obstruent, do not operate across word boundaries, their lack of activity provides a crucial clue that word-level morphology has taken place. The data indicate that for speakers who are particularly sensitive to phonological information (see Thomas *et al.*, 2001), and who have an intact rule system, morphological paradigms are easier to learn when illegal phonotactics are created.

One puzzle is why the typically developing children tested here do not generally show the same advantage for illegal verbs. Recall that the oldest group in van der Lely and Ullman's experiment did, but the other two groups in that study did not, and in Thomas *et al.*'s experiment the numerical advantage for illegal verbs was not significant. Given extensive research that babies aged six months and older are able to extract phonotactic regularities from speech (for a review, see Jusczyk, 1997), it seems plausible that the illegal clusters that signify a morphological boundary could be used as a parsing aid in learning verb paradigms. It may be that this gives children an advantage for regular verbs containing illegal clusters at a younger age than has been investigated here, but that once a symbolic system for morphology is in place, phonotactic cues no longer play a role.

The phonotactic characteristics of the verb-end cluster clearly affect the performance of the G-SLI group on regular past tense morphology. In line with van der Lely and Ullman's claim, the results can be explained by proposing that G-SLI children have an impairment in the past tense suffixation rule. For van der Lely and Ullman, this rule is not missing from the grammar: it operates optionally, and so unlike normally developing children, G-SLI children have to rely on the storage of past tense forms.

Another possibility, which van der Lely and Ullman do not discuss, but which van der Lely (1997b) does, is that G-SLI children are able to form past tense regular forms through analogy with already known past tense forms and monomorphemic words with the same ending. Legality could therefore play a role twice: (1) legal past tense forms could be stored more effectively than illegal ones, as a result of their higher frequency, and (2) when creating past tense forms by analogy, verbs with legal clusters will be easier to create because of the higher frequency of legal clusters. Because irregular verbs do not contain illegal clusters, these findings suggest that phonotactic illegality could be a contributing factor to the disproportionate difficulty that G-SLI children have with regular as opposed to irregular morphology.

Chapter 4. Confirming the morphological defcicit: Further exploring the impact of verb-end phonotactics

4.1. Introduction

4.1.1. Chapter outline

The work reported in Chapter 3 revealed that in an elicitation task G-SLI children have more difficulty with verbs whose inflection creates a phonotactically illegal cluster than with verbs whose cluster is legal. I interpreted this as indicating that children with G-SLI have difficulty creating morphologically complex forms. However, the data in Chapter 3 were analysed post hoc and contained only a small set of verbs. A further concern is that the performance of the G-SLI group was much lower than that of the typically developing controls. It is possible that G-SLI children show the same pattern of performance as younger typically developing children would, and that we cannot tell whether this is the case because the youngest control group was not young enough (Michael Thomas, personal communication, September 2003). The aim of this chapter is to confirm the morphological deficit in G-SLI by carrying out a task which elicits the past tense form in a way that is intended to increase inflection levels, and which uses a larger number of verbs.

4.2. Method

4.2.1. Verb stimuli

Two conditions, with 8 verbs in each condition, were selected. One condition contains verbs which have a legal cluster at the inflected verb-end, and the other contains verbs with an illegal cluster at the verb-end. All verbs have a two-consonant cluster in the inflected form, in order to maintain constant prosodic complexity. It proved impossible to balance the conditions for past tense frequency given the constraints on which verbs could be used – not only did the verbs have to be familiar to children of the ages taking part in the experiment, but they had to enter into stimulus sentences that were syntactically correct and pragmatically plausible. However, if a correlation is found between frequency and performance, then frequency can be partialled out of the analysis. Table 4.1 shows the characteristics of the stimuli. For the full list of verbs, see the Appendix A.1.

Frequency values are the raw frequencies augmented by 1 and ln-transformed. Two different measures are used: CO-BUILD, and Francis and Kucera. The CO-BUILD measure is used because this is the one used in van der Lely and Ullman's (2001) study, and because it is available in an electronic format which makes the calculation of cluster frequencies possible. The Francis and Kucera measure is used because this is the database I use for the experiments in future chapters.

Which frequency measures one uses depends on how one thinks children produce past tense forms. If one assumes that they create regular past tense forms by rule, then the frequency measure to use is the sum of the frequencies of the base form and all its inflectional variants, i.e. for *played* this would be *play*, *plays*, *played* and *playing*. If one assumes that the past tense form is stored, then the relevant frequency measure is just the frequency of the past tense form (see Alegre & Gordon, 1999). Obviously, van der Lely and Ullman's model is one whereby typically developing children use a rule for regulars (and therefore the full paradigm frequency is appropriate), while the G-SLI children store regulars, and all children store irregulars (and so the past tense frequency is appropriate). Being faced with having to choose one or the other, I decided for this experiment to use the frequency of the past tense form only.

Table 4.1. Characteristics of the stimuli

Condition[#]	**Examples**	**Past tense frequency**		**Cluster frequency**
		CO-B*	**F&K****	**CO-B***
VC-D legal	*killed, wrapped*	1.647	2.089	7.726
VC-D illegal	*touched, robbed*	1.175	1.425	4.785

[#] -D is the symbol used to indicate the coronal stop, whose voicing quality is determined by the voicing of the preceding consonant. * Frequencies obtained from CO-BUILD, CELEX database; ** Frequencies obtained from Francis & Kucera

4.2.2. Procedure: Elicitation task

The procedure was based on that used by van der Lely and Ullman (2001). In their task the lead in was of the form:- '*Everyday I rob a bank. Yesterday, just like everyday, I ______ a bank*'. The format used here has been changed in an attempt to raise the level of correct performance for the G-SLI children, who achieved a score of only 22.2% (for regulars) in van der Lely and Ullman's task, and many of whom perform poorly on the Verb Agreement and Tense Task (VATT, see Section 2.2). The lead in for the task reported in this chapter presents both the past tense and bare stem form:- '*Last week Kipper robbed a post office. Everyday I rob a post office. Yesterday I ______*'. Hence the lead in primes the syntactic form of the past tense, as does the inclusion of only regular stimuli. By raising the level of performance it was hoped to get a wider spread of scores in the different verb groups, hence avoiding any possible floor effects.

The experimenter introduces the child to two toy dogs. One of the dogs, Kipper, is likely to be familiar to the child through the popular children's books and television programmes. The experimenter reminds/tells the child that Kipper is a very adventurous dog which does all sorts of exciting things. The second dog, which the child won't know, is called Bean Dog. The experimenter tells the child that Bean Dog is one of Kipper's best friends, but that he gets very jealous of Kipper and all the adventures that Kipper has. The experimenter explains that Kipper has been busy doing lots of things recently. Bean Dog is jealous and wants to tell everyone that he has been doing them too. The experimenter asks the child to help her to be the voice of Bean Dog and tell everyone the things that he has been doing. There are 4 practice items using irregular verbs, and 16 experimental items which are listed in Appendix A.2. One pseudo-randomised list was created for all participants.

4.2.4. Participants

14 G-SLI children participated, and 28 typically developing children were selected as controls. 14 of the control children were matched on raw score (to within ±1 point) on a test of sentence comprehension, the Test for Reception of Grammar (TROG; Bishop, 1983) and 14 matched on raw score (±3) on the British Picture Vocabulary Scales (BPVS; Dunn *et al.*, 1997). These children were then divided into two groups according to age, in order to give a picture of typical development. The Language Ability 1 (LA1) control group are aged 4;06-7;05, with a mean age of 6;00, and the Language Ability 2 (LA2) control group are aged 7;06-12;00, with a mean age of 9;06. Details of the participants are presented in Table 4.2.

Table 4.2. Participant details

Measure		G-SLI N = 14	LA1 controls N = 14	LA2 controls N = 14
Age	**Mean**	12;03	6;00	9;06
	Range	9;09 – 16;08	4;06 – 7;05	7;06 – 12;00
TROG	**Raw, mean**	12.86	10.76	16.43
	Raw, range	6 – 17	6 – 16	12 – 19
	z-score, mean	-1.67	-0.14	0.12
BPVS	**Raw, mean**	79.93	60.00	94.21
	Raw, range	47 – 104	33 – 81	69 – 120
	z-score, mean	-1.67	0.28	0.28

A series of independent samples t-tests revealed no significant difference between the G-SLI group and the LA1 control group on the TROG, t (26) = 1.062, p = 0.298, but a significant difference between these two groups on the BPVS, t (26) = 3.172, p = 0.004. The difference between the G-SLI and LA2 groups was significant for both the TROG and the BPVS, t (26) = -3.364, p = 0.002 and t (26) = -2.243, p = 0.034, respectively. The LA1 control group therefore provides a grammar age match for the G-SLI group. In terms of vocabulary ability, the G-SLI group falls between the LA1 and LA2 groups.

4.2.5. Predictions

Predictions for the elicitation task are different for the G-SLI and the control groups. For the G-SLI group I predict that accuracy will be lower for verbs with illegal clusters compared to those with legal clusters. For typically developing children I predict no effect of cluster phonotactics, although if there is an effect I expect higher levels of accuracy on verbs with illegal clusters compared to those with legal clusters, i.e. the opposite direction to the G-SLI group. The majority of errors for all groups are predicted to be bare stem errors.

4.2.6. Coding of responses

Responses were coded as follows:-

- Correct — correct inflection, e.g. *hop* → *hopped*
- Bare stem — inflection missing, e.g. *rob* → *rob*
- Other responses — e.g. *wash* → *washing*; *judge* → *jumped*

On the rare occasions when a child corrected himself, the first response was accepted for analysis.

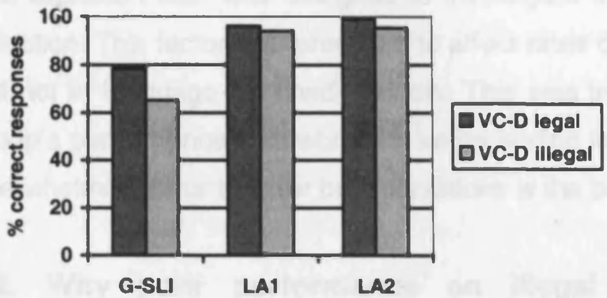
4.3. Results

Data from one participant in the LA1 group were discarded because her answers were at times muffled and therefore difficult to transcribe accurately. Correct responses to the elicitation task are shown in Table 4.3 and illustrated in Figure 4.1.

Table 4.3. % correct responses

Condition		G-SLI (N=14)	LA1 (N=13)	LA2 (N=14)
VC-D legal	**Mean (SD)**	78.57 (23.22)	96.15 (7.93)	99.12 (3.34)
VC-D illegal	**Mean (SD)**	65.18 (38.97)	94.23 (10.96)	95.54 (10.52)

Figure 4.2. % correct responses according to phonotactics



Note that a correlation between correct performance and verb frequency (both measures) revealed no significant or even marginally significant effects of frequency on the performance of any of the three participant groups. Therefore frequency is not used as a covariable in any of the analyses performed on these data.

Correct performance was analysed using a 3 (Group: G-SLI, LA1, LA2) x 2 (Condition: VC-D legal, VC-D illegal) ANOVA. This revealed significant main effects of group, $F(2, 38) = 7.745$, $p = 0.002$, and condition, $F(1, 38) = 9.079$, $p = 0.005$. The interaction between group and condition was marginally significant, $F(2, 38) = 2.957$, $p = 0.064$. T-tests showed that for the G-SLI group, performance on VC-D legal verbs was significantly higher than performance on VC-D illegal verbs, $t(13) = 2.446$, $p = 0.029$. For the LA1 and LA2 groups, however, the pairwise comparisons between the two conditions did not reach significance, $t(12) = 1.000$, $p = 0.337$, and $t(13) = 1.749$, $p = 0.104$ respectively. The pattern of correct performance is therefore VC-D legal > VC-D illegal for the G-SLI group, but VC-D legal = VC-D illegal for the controls.

The majority of errors for the all groups comprise bare stem responses. The proportion of these errors expressed as a percentage of the total number of responses is shown in Table 4.3.

Table 4.3. % Bare stem errors

Stimulus	G-SLI (N = 14)	LA1 (N = 13)	LA2 (N = 14)
VC-D legal	16.96 (18.74)	3.85 (7.88)	0.89 (3.34)
VC-D illegal	26.79 (32.84)	4.81 (10.87)	4.46 (10.52)

4. Discussion

4.1. Summary of results

The elicitation task was designed to investigate the impact of verb-end phonotactics on inflection. This factor was predicted to affect rates of past tense inflection in G-SLI children but not in language-matched controls. This was indeed found to be the case: the G-SLI group's performance was worse on verbs ending in illegal compared to legal clusters. The overwhelming error type for both conditions is the bare stem form.

4.2. Why poor performance on illegal verbs is *not* indicative of a phonological deficit in G-SLI

The results presented in this chapter confirm my analysis in Chapter 3, that there is a morphological deficit in G-SLI. It is worth clarifying at this stage why I do not interpret the disadvantage for illegal forms as indicating a phonological deficit. It is of course possible that G-SLI children find marked sequences of segments more difficult than unmarked sequences. For example, they might have difficulties with nasal-obstruent clusters which do not agree in pace of articulation, e.g. */md/*, and therefore be more successful at inflecting verbs when the cluster shares place features, e.g. */nd/*. This hypothesis is difficult to test precisely because segmentally marked sequences occur across the stem-affix boundary, making it impossible to tease apart a phonological effect from a morphological one. However, I discount the possibility that the locus of this particular deficit is in the phonology for one crucial reason: the perceptual deficit that has been proposed by Tallal, Joanisse, Seidenberg and others makes, as far as I can tell, no predictions as to which particular sequences of segments would be more susceptible to impairment than others. I can think of no perceptual reason why the suffix should be less likely to be perceived after a heterorganic nasal than a homorganic one, or after a voiced obstruent rather than an unvoiced one. I therefore claim that the pattern of performance on regular verbs with illegal versus legal clusters is evidence of a morphological deficit in G-SLI, and therefore of a dual mechanism model of inflection in typical development

The dual versus single mechanism debate has also been raging in the literature on adult aphasia (e.g. Bird, Lambon Ralph, Seidenberg, McClelland & Patterson 2003; Tyler, Randall & Marslen-Wilson, 2002). Double dissociations between regular and irregular past tense performance have been reported (e.g. Ullman, *et al.*, 1997; see Marslen-Wilson & Tyler, 1997, for double dissociations in aphasic patients), but proponents of both models claim to be able to account for this finding. I will pick up just two pertinent points in the discussion here, both of which relate to Bird *et al.*'s (2003) paper. Bird *et al.* studied 10 patients with non-fluent aphasia, who suffer an apparent disadvantage for regular past tense forms. The authors found that differences between regulars and irregulars were not significant once the stimuli were controlled for phonological complexity, and claimed that a single mechanism model could explain this finding. However, this result is exactly what one would expect if aphasic individuals are using a single mechanism – as Ullman *et al.* claim, and as van der Lely and Ullman (2001) claim for the G-SLI group! In some ways the data from the aphasic and G-SLI populations are a red herring – the dual and single mechanism accounts make same prediction for these groups, and so cannot be used to tease the two models apart. Instead the models need to be able to explain the typically developing data. This requirement for explaining the pattern of typical behaviour leads me to my second point. It seems that the single mechanism model has more difficulty trying to explain the regularity advantage for the typically developing children in van der Lely and Ullman's study. A single mechanism model would surely have to predict that regulars with illegal clusters would be more difficult for typically developing children too – but the evidence presented in this and the previous chapter is that they are not.

Returning to van der Lely and Ullman's (2001) study: the authors found a very striking impairment for both regulars and irregulars in the G-SLI group. However, they barely address the hypothesis that past tense inflection in G-SLI could be impaired as a result of two separate deficits in different components of the grammar: one in morphology and one in syntax. Van der Lely has proposed that that the syntactic deficit in G-SLI can be characterised by optional syntactic Movement for feature checking. This hypothesis, the Representational Deficit for Dependent Relations (RDDR), can partially account for the G-SLI group's poor performance on this past tense task, because feature checking for Tense involves movement from V to I (see Section 2.1.3). If Tense features do not move, they cannot be checked, and therefore not spelt out on the verb, resulting in a bare stem form (see Davies, 2001). So the RDDR can account for low levels of past tense marking on both sets of verbs. The reason that performance on regulars is even lower than expected is due to a separate deficit in past tense suffixation, as proposed in van der Lely and Ullman. Following the proposal that deficits in syntactic feature checking and morphological rule-use impact on tense marking, I hypothesise that if deficits exist in a further module of grammar – phonology – then those will impact on regular inflection too, given that regular past tense verb-ends frequently contain clusters. This hypothesis will be tested in the next four chapters. In Chapter 5 I show that most children with G-SLI also have a deficit in complex phonological representations, and in Chapters 6, 7 and 8 I show how this deficit impacts on their past tense inflection.

Chapter 5. Establishing the phonological deficit – the impact of prosodic complexity on phonological representations

5.1. Introduction

5.1.1. Chapter outline

The aim of this chapter is to characterise the phonological abilities of G-SLI children using a non-word repetition test (Test of Phonological Structure, van der Lely & Harris, 1999) that systematically varies syllabic and metrical structure. In Section 5.1.2 I discuss previous studies of non-word repetition in SLI, and in Section 5.1.3 I explain how the TOPhS enables us to characterise the phonology of SLI children. In Section 5.2 I present the method, and in Section 5.3 the results. Most G-SLI children perform poorly on the TOPhS. However, there is a wide range of abilities within the G-SLI group, so in Section 5.3.2 I present a more detailed analysis of the performance of three G-SLI children chosen to illustrate this range. I also present data from a child (who was tested as a candidate for the control group) who has poor phonology but normal grammatical abilities. In Section 5.4 I propose an Optimality-Theoretic account of data from one particular G-SLI child, whose onset clusters demonstrate positional markedness effects. I show that typically developing children and some of the other children in the G-SLI group also show these effects, and I argue that this phenomenon is one of the ways in which metrical structure and syllable complexity can interact in influencing the shape of children's phonological outputs. In Section 5.5.1 I propose a model of prosodic representations in children with G-SLI, and in Section 5.5.2 I discuss how the data militate against a causal relationship between poor phonology and a grammatical deficit.

The group data have previously been presented in Marshall, Harris and van der Lely (2003), and data from one of the children, GD, are discussed in Marshall, Ebbels, Harris and van der Lely (2002).

5.1.2. Non-word repetition abilities in SLI

Non-word repetition tasks have been used by many researchers to investigate the phonological abilities of children with SLI. The most widely used is the Children's test of Non-word Repetition (CNRep; Gathercole & Baddeley, 1996), which consists of forty non-words between two and five syllables long. These non-words are presented either on cassette tape or by the administrator, and the child repeats them immediately. Each repetition attempt is scored as either correct or incorrect. The CNRep appears to be a robust clinical marker for SLI, with children's performance deteriorating as syllable number increases (e.g. Bishop, North & Donlan, 1996; Dollaghan & Campbell, 1998; Gathercole & Baddelely, 1990). However, poor CNRep scores are not confined to the SLI population. Children with Down's Syndrome (Jarrold, Baddeley & Hewes, 2000), autism (Kjelgaaard & Tager-Flusberg, 2001) and dyslexia (Ramus, Rosen, Dakin, Day, Castellote, White & Frith, 2003) also perform badly on this test.

Gathercole and Baddeley (1990) propose that the CNRep taps into children's phonological short-term memory abilities. They claim that SLI children perform poorly on the test because they have limited capacity in their phonological store, and/or an unusually rapid decay rate for items held there. If children are poor at retaining a short-term representation of speech sounds, they are likely to have difficulty in forming long-term representations of new words. This in turn impacts on the identification of syntactic structures, because word sequences are not retained long enough for grammatical analysis. In other words, Gathercole and Baddeley propose that phonological short-term memory deficits are primary in SLI, and that other language problems arise as a consequence.

Gathercole and Baddeley's claims have not gone unchallenged, however, and alternative explanations that take a psycholinguistic perspective have been proposed. Snowling, Chiat and Hulme (1991) stress that the difference between phonological memory and other phonological processes, such as phonological segmentation and articulatory execution, cannot be ignored when interpreting the results of the CNRep. Van der Lely and Howard (1993) argue that the causal arrow is reversed, so that linguistic deficits are actually the cause of phonological short-term memory deficits. In a similar vein, Edwards and Lahey (1998) hypothesise that the deficit lies not in the ability to hold phonological information in short-term memory, but rather in the formation or storage of phonological representations.

The effectiveness of the CNRep as a measure of phonological short term memory relies on the assumption that the child has no lexical representation for the particular sound pattern he is asked to repeat. Unfortunately this assumption is not met, since many items contain real words within them, including *hampent*, *defermification*, *underbrantuand* and *reutterpation*. My colleagues and I have claimed that it is easier to create a phonological representation of a non-word when a portion of it can be retrieved from long-term memory, so that the entire non-word does not need to be created *de novo* (Marshall *et al.*, 2002). A similar point can be made about inflectional and derivational morphemes. Some of Gathercole and Baddeley's non-words have suffixed endings, as in *blonterstaping*, *defermication*, *loddernapish*, *contramponist*. It follows that children with large vocabularies and/or a good knowledge of morphological structure are more likely to make analogies with familiar words, thereby gaining higher scores. As children with SLI tend to have poorer vocabularies and impaired morphological abilities, such deficits could account for, or at least contribute to, poor performance on the CNRep. A correlation between poor CNRep scores and SLI is therefore not surprising.

In support of our view that lexical factors can account for poor non-word repetition in SLI are studies showing that lexical factors do influence non-word repetition in typically developing children. One such factor is word-likeness: Dollaghan, Biber and Campbell (1995) found that children repeat non-words with stressed syllables that correspond to real words significantly more accurately than those with stressed syllables that are non-lexical. If these words are familiar to the child, Dollaghan *et al.* claim that capacity is freed up in working memory for remembering a greater number of syllables. Likewise, syllable frequency affects non-word repetition performance, with accuracy being higher on non-words containing syllables that are more frequent in English polysyllabic words (Nimmo & Roodenrys, 2002).

There are also phonological concerns over the non-words chosen for the CNRep. Syllable number is the only variable along which children's performance is measured. Yet within a set of non-words of identical syllable number, various types of syllable and foot structure occur. Compare *pennel* and *glistow* – the first has no consonant clusters whereas the second has two (*gl, st*). Syllabic complexity has been claimed to influence non-word repetition. Gathercole and Baddeley (1990) found that non-words with consonant clusters were harder for children to repeat, although the effect was similar for both typically developing and language-impaired participants. They interpreted this difficulty with clusters as being related to articulation problems. In contrast, Bishop, North and Donlan (1996) found that while consonant clusters affected repetition accuracy in both groups, the effect was significantly greater for the SLI group. Now compare the non-words *blonterstaping* and *perplisteronk* – the first consists of two trochaic feet (*blonter* and *staping*), whereas the second consists of two trochees (*plister* and *onk*) and an initial unfooted syllable (*per*). Initial weak syllables are known to cause difficulties for SLI children (Aguilar-Mediavilla, Sanz-Torrent & Serra-Raventos, 2002; Sahlen, Reuterskiold-Wagner, Nettelbladt & Radeborg, 1999).

The design of the CNRep does not allow a fine-grained investigation of which particular prosodic structures cause the most errors. So while performance might indeed *correlate* with language abilities, it is unjustified to conclude that non-word repetition difficulties are *caused* by an increase in syllable number, and by extension a deficit in phonological short-term memory. Correlation is not the same as cause, and the deficit might instead be in forming correct phonological representations in the first place rather than in retaining them.

5.1.3. Characterising the phonology of G-SLI children

The Test of Phonological Structure (TOPhS, van der Lely & Harris, 1999) tests the effects of prosodic complexity on non-word repetition performance. Stimuli are varied according to the number of marked syllabic and metrical structures they contain (see Section 1.2.3.2 for a discussion of markedness). Three of the marked structures relate to syllable structure and two to metrical structure. The three syllabic structures are set out in Table 5.1, together with real words and examples drawn from the non-word stimulus set. In each of the examples, the segment string illustrating the relevant structure is underlined. Marked and unmarked syllabic structures are compared only in the stressed syllable.

Table 5.1. Syllabic structures varied in the TOPhS

	Syllabic structure		**Real word**	**Non-word**
Onset	no cluster	unmarked	*pawn*	*ketə*
	cluster	marked	*prawn*	*kletə*
Rhyme	open	unmarked	*city*	*ketə*
	closed	marked	*filter*	*kestə*
Word end	V-final	unmarked	*city*	*ketə*
	C-final	marked	*sit*	*ket*

As to metrical structure, the TOPhS is designed to vary the location of the stress foot relative to word edges. In the unmarked case, the edge of a foot is aligned with the edge of the word. Words consisting of a single foot have perfect alignment at both edges (as in *city, tea, sit*). In polysyllabic words, misalignment is possible, resulting in marked stress patterns. Two of these feature in the non-word stimulus set, both involving the adjunction of an unstressed syllable at a word's edge. In one pattern, an unfooted syllable is adjoined at the beginning of a word, as in *ba{nana}, de{nial}* (feet parenthesised). The other involves right-edge adjunction, where an unfooted syllable separates the end of a foot from the end of a word, resulting in antepenultimate stress, as in *{Jenni}fer, {fanta}sy*. The structures are summarised and exemplified in Table 5.2.

Table 5.2. Metrical structures varied in the TOPhS

Metrical structure		Real word	Non-word
Left adjunction	unmarked	*{city}*	***{ketə}***
	marked	*ba{nana}*	***fə{ketə}***
Right adjunction	unmarked	*{city}*	***{ketə}***
	marked	*{Cana}da*	***{ketə}lə***

In certain respects, string-based and prosody-based measures of complexity converge. For example, the extra segment that renders *play* longer than *pay* also contributes to the complexity of the onset in *play*. In other respects, however, the two types of measure produce quite different results. For example, on a phoneme or syllable count, *tidy* and *today* are of equal complexity. However, in terms of metrical structure, *today* is more complex than *tidy* by virtue of containing a left-adjoined syllable.

The TOPhS requires the child to repeat non-words that are systematically varied with respect to the five marked structures described above. The stimuli were constructed around four exemplars of CVCV structure which are manipulated to contain 24 different permutations of marked and unmarked structures, yielding a total of 96 stimuli. Each set contains stimuli ranging from a maximally simplex form, displaying only unmarked structures (e.g. ***ketə***), through progressively more complex forms, containing various permutations of marked structures (e.g. ***fəkestələ***). Table 5.3 provides illustrative examples of non-words based on the CVCV form ***depə***, where, 'u' and 'm' indicate unmarked and marked structures respectively. All non-words conform to the phonotactic constraints of English and are intended to be applicable to all dialects of English. In this way, the TOPhS allows us to test the prediction that non-words with marked structures will be repeated less accurately than those with unmarked structures.

Table 5.3. Examples of non-words based on the CVCV form *depə*

non-word	onset	rhyme	word-end	left adjunction	right adjunction
depə	u	u	u	u	u
drepə	m	u	u	u	u
dempə	u	m	u	u	u
dep	u	u	m	u	u
bədepə	u	u	u	m	u
depəri	u	u	u	u	m
bədrepə	m	u	u	m	u
dempəri	u	m	u	u	m
bədrempəri	m	m	u	m	m

5.1. Method

5.2.1. Procedure

Testing was carried out in a quiet room in the children's school. The children were told that they were going to hear some made-up words that they would not have heard before and that they should repeat these words into the microphone. They listened to the digitally recorded non-words through high quality headphones and their repetitions were recorded onto a DAT tape. Four practice items were provided at the start of the task, and the 96 non-words were then presented in a set randomised order, at three second intervals. The time taken to complete the task was approximately 6 minutes.

The participants' repetitions were transcribed online in broad phonetic IPA transcription and then subsequently verified against the recording. For the purposes of the statistical analyses responses were scored as either correct or incorrect. Voicing errors, e.g. /*p*/ for /*b*/ were not scored as incorrect, and neither was replacement of /*r*/ by /*w*/.

5.2.2. Participants

10 G-SLI children were administered the TOPhS. In addition, 20 children with typically developing language acted as controls. 10 children were individually matched to the G-SLI participants on exact raw score on the Test for Reception of Grammar (TROG; Bishop, 1983) (with the exception of one G-SLI child, GS, whose grammar control's score is 3

points above). 10 children were individually matched on raw score (±3) on the British Picture Vocabulary Scales (BPVS; Dunn *et al.*, 1997). In order to get a picture of typical development, those 20 children were then divided into two groups according to age. Details of the participant groups are shown in Table 5.4.

Table 5.4. Participant details

Measure		G-SLI N = 10	LA1 controls N = 10	LA2 controls N = 10
Age	Mean	12;00	5;09	8;09
	Range	9;04 – 16;08	4;05 – 7;04	7;05 – 9;10
TROG	Raw, mean	13	10.6	16.4
	Raw, range	6 – 17	6 – 15	12 – 19
	z-score, mean	-1.60	-0.20	0.31
BPVS	Raw, mean	76.5	58.9	89.6
	Raw, range	47 – 104	33 – 80	69 – 102
	z-score, mean	-1.67	0.13	0.21

T-tests reveal that the G-SLI group does not score significantly differently to either the LA1 or LA2 group on the TROG, $t(9) = 1.933$, $p = 0.085$, and $t(9) = -2.037$, $p = 0.072$ respectively. The LA1 group is hence the better match in terms of grammar ability. The G-SLI group scores significantly better than the LA1 group on the BPVS, $t(9) = 2.258$, $p = 0.050$, but not significantly differently to the LA2 group, $t(9) = -1.737$, $p = 0.116$. The LA2 group is therefore the better match in terms of vocabulary ability.

5.3. Results

5.3.1. Group results

The scores for the G-SLI group and the LA control groups are set out in Table 5.5.

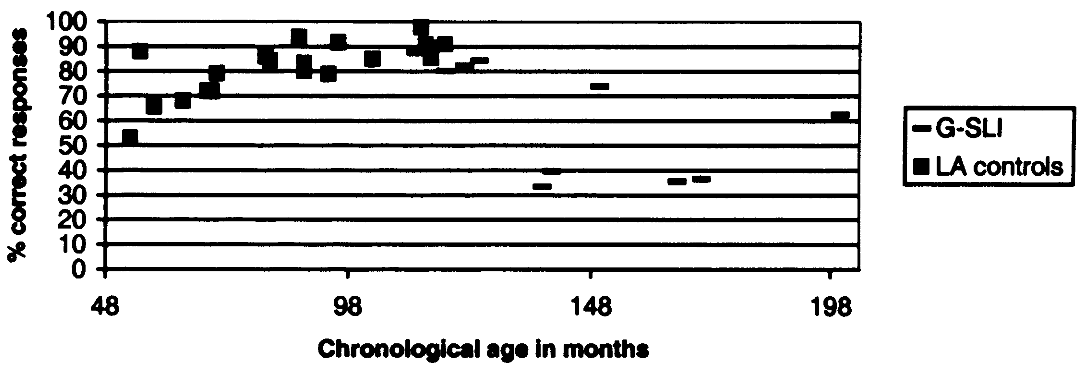
Table 5.5. % correct scores on the TOPhS

	G-SLI	LA1	LA2
Mean score (SD)	61.56 (22.90)	77.40 (11.94)	87.19 (6.23)
Range of scores	33.33 – 87.50	55.21 – 91.76	79.17 – 97.92

A one-way ANOVA on correct scores revealed a significant effect of group, $F(2, 29) = 7.108$, $p = 0.003$. Post hoc comparisons (Bonferrorni-corrected) revealed that the G-SLI group did not perform significantly differently to the LA1 group, $p = 0.087$ but did perform significantly worse than the LA2 group, $p = 0.003$. The difference between the two control groups was not signifcant, $p = 0.495$.

There is a wide range of performance within the G-SLI subgroup. There is also a wide variety of ages in the G-SLI and control groups. In order to obtain a picture of how performance changes with age, score is plotted against age for the G-SLI and LA controls in Figure 5.1.

Figure 5.1. TOPhS score plotted against age



For the LA children as a whole, performance on the TOPhS is positively correlated with age, $r = 0.728$, $p < 0.001$. For the G-SLI group, performance on the TOPhS is *not* correlated with age, $r = -0.535$, $p = 0.110$. This absence of a correlation in the G-SLI group is due to 4 older children performing below the level of even the 5-year old control children, whereas 4 of the younger G-SLI children perform near chronological age-appropriately. We might therefore ask whether other language measures correlate with age in the G-SLI group, and indeed they do: age is strongly positively correlated with scores on both the TROG, $r = 0.844$, $p = 0.002$, and the BPVS, $r = 0.873$, $p < 0.001$. In other words, whereas the grammatical and vocabulary abilities of G-SLI children improve over time, phonological abilities do not necessarily do so.

For the LA children, TOPhS score correlates with BPVS score, $r = 0.462$, $p = 0.040$ but not with TROG score, $r = 0.151$, $p = 0.524$. For the G-SLI group however, TOPhS score correlates with neither BPVS nor TROG scores, $r = -0.559$, $p = 0.094$, and $r = 0.151$, $p = 0.524$, respectively. In other words, for the G-SLI group phonological abilities as measured by the TOPhS do not correlate with measures of receptive language.

5.3.2. Individual results

The whole group analysis conceals a wide range of individual scores within the G-SLI group, and it is therefore worth looking at three children - GD, LJ and BD - who exemplify this range. The question of interest is how these children's performance on the TOPhS compares with that of their language-age matches. In order to provide large enough numbers of the appropriate language-age controls, I selected further typically developing children according to the criteria outlined in Section 5.2.2.

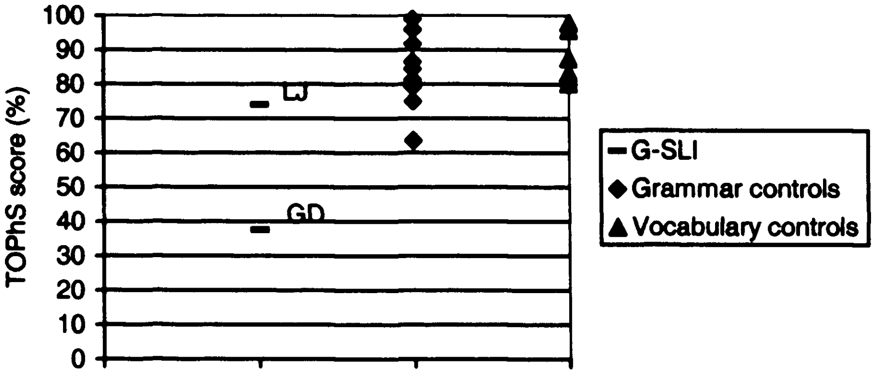
5.3.2.1. Children GD and LJ

Because GD and LJ have identical scores on the BPVS and such similar scores on the TROG, the same language matches are used for each. Grammar controls have raw scores of 16 and 17 as measured by the TROG, and vocabulary controls have raw scores of 92±3 as measured by the BPVS. There are 12 grammar controls and 7 vocabulary controls. The difference in these numbers reflects the fact that the BPVS has a much wider range of scores than the TROG, and therefore it is harder to find children with a particular score. Details for GD and LJ, and their grammar and vocabulary matched controls, are presented in Table 5.6.

Table 5.6. TOPhS scores: GD, LJ and language controls

		GD	LJ	Grammar controls (N = 12)	Vocabulary controls (N = 7)
TROG raw	**Mean (SD)**	17	16	16.33 (0.49)	n/a
BPVS raw	**Mean (SD)**	92	92	n/a	91.86 (1.86)
TOPhS %	**Mean (SD)**	37.50	73.96	84.98 (9.91)	87.20 (7.13)
Age	**Mean (SD)**	14;03	12;06	9;00 (1;07)	8;06 (0;09)

Figure 5.2. TOPhS scores: GD, LJ, and language controls



GD's score is well outside the range of scores found in the grammar control group, and is 4.79 SD below the mean. GD's score is also well outside the range of scores found in the vocabulary control group, and is 6.96 SD below the mean. LJ's score, however, is within range of the grammar controls and is only 1.11 SD below mean. His score is just outside the range of the vocabulary controls and is 1.86 SD below the mean. We can conclude on the basis of this analysis that one G-SLI child, LJ, has near normal performance on the TOPhS relative to his language age, and one G-SLI child, GD, scores well below what would be expected given his language age. This is despite the fact that GD and LJ have identical BPVS and near-identical TROG scores.

5.3.2.2. Child BD

BD's TOPhS score of 87.5% indicates that his performance may be above that of his language-matched peers. Table 5.7 shows that relative to his two language-matched controls selected for the analysis in Section 5.3.1, this is indeed the case.

Table 5.7. TOPhS scores: BD and language controls

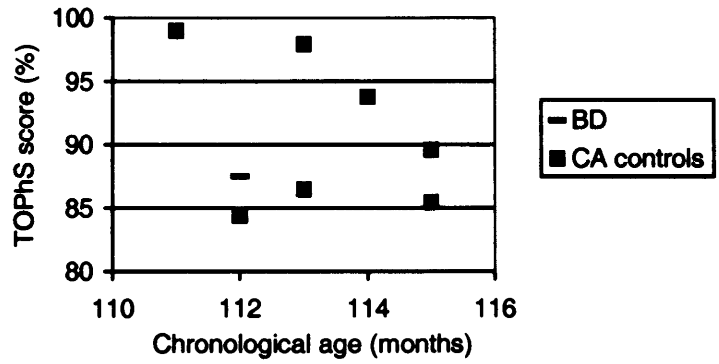
	GD	Grammar control	Vocabulary control
TROG raw	6	6	n/a
BPVS raw	47	n/a	44
TOPhS %	87.5	65.63	55.21
Age	9;04	4;10	4;05

The confirmation that BD's TOPhS score is higher than that expected based on his language scores raises the possibility that his performance is actually chronological age-appropriate. I therefore compare his score with those of eight chronological age-matched controls, aged within ±3 months. These details are shown in Table 5.8 and Figure 5.3.

Table 5.8. TOPhS scores: BD and chronological age-matched (CA) controls

	GD	CA controls (N = 8)
TOPhS %	87.5	91.80 (6.08)
Age	9;04	9;05 (0;01)

Figure 5.3. TOPhS scores: BD and CA controls



BD's TOPhS score is within 1SD of that of his chronological age-matched controls. We can conclude that not every G-SLI child has poor phonology.

5.3.2.3. Child WD

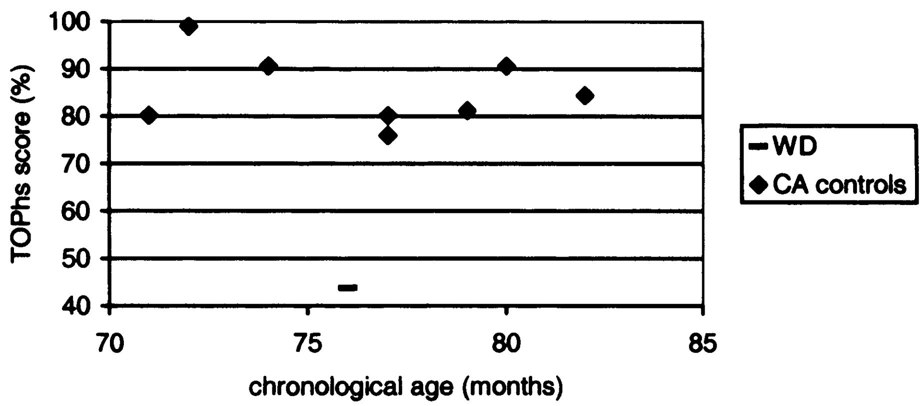
Is it possible for a child to show the reverse pattern, i.e. to perform poorly on the TOPhS and yet still have intact grammar and vocabulary skills? Data from WD provide evidence that phonological impairments do not inevitably lead to poor language skills. WD was tested as a potential control participant. She has normal receptive grammar: her standard score as measured by the TROG is 87, which is in the low normal range but still not low enough to be indicative of a language deficit. Her receptive vocabulary as measured by the BPVS is also normal, with a standard score of 101. Her non-verbal IQ, as measured by the BAS, is within the normal range (standard score 92.5).

When WD's TOPhS score is compared to those of 9 children of the same chronological age (±6 months) it can be seen that her score is far below what one would expect for her age (6.24 SD). These details are shown in Table 5.9 and Figure 5.4. It can be seen that although her performance on standardised tests of receptive language is age-appropriate, her phonological skills are not.

Table 5.9. TOPhS scores: WD and CA controls

	WD	**CA controls (N = 9)**
TOPhS %	43.75	86.23 (6.81)
Age	6;04	6;05 (0;04)

Figure 5.4. TOPhS scores: WD and CA controls



5.3.3. Group analysis of performance according to phonology

Previous researchers have analysed the results of the TOPhS in different ways. Ebbels used multiple regression, with onset, rhyme, left adjunction, right adjunction and syllable number as predictor variables (Ebbels, 2003). Word end is missing from this list because it is not independent from right adjunction – in the TOPhS stimulus set, a non-word with a right adjoined syllable is never marked for word end (i.e. does not end in a consonant), and therefore the multiple regression procedure removes it from the analysis. Ebbels finds that onset, rhyme and left adjunction are all significant predictors of performance, with left adjunction being the most significant. Syllable number is not a significant predictor of performance

Gallon, Harris and van der Lely (submitted) used ANOVA to analyse their data. This is problematic because this method assumes that factors are independent (Field, 2000), and in the current design of the TOPhS word end and right adjunction are not. There is also the problem that the two metrical predictors, left and right adjunction, are not independent of syllable number, as they both add a syllable to the non-word. Syllable number is not included as a factor in Gallon *et al.*'s analysis, presumably because of this lack of independence.

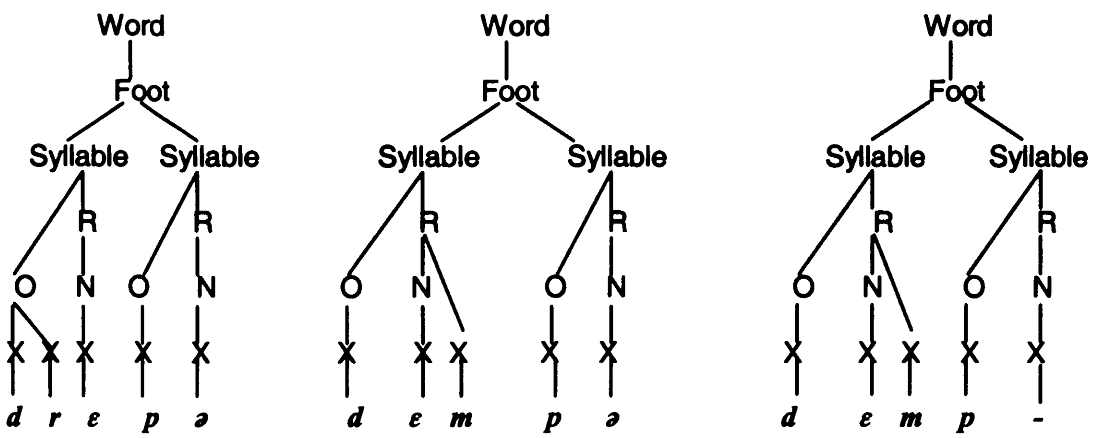
I have chosen to carry out the analysis of my data a little differently. The analysis in Section 5.3.3.1 looks at syllable structure by considering how many consonant clusters are present in the non-word. The analysis in Section 5.3.3.2 looks within two-syllable, and then three-syllable, non-words to compare the effect of contrasting metrical structure independent of syllable number.

5.3.3.1. Analysis according to syllabic complexity and syllable number

Consonant clusters can arise in three ways, and their structure is shown in (i) – (iii) in Figure 5.5. A maximum of two clusters can be present in a non-word in the TOPhS stimulus set, as cluster types (ii) and (iii) are clearly mutually exclusive: (ii) is always word-medial and (iii) is always word-final. (i) can occur word-initially or following an initial weak syllable (e.g. *bədrεpə*).

Figure 5.5. Prosodic structures which contain consonant clusters

(i) Marked onset (ii) Marked rhyme (iii) Marked rhyme and word end



In Section 5.3.3.1.1 I present an analysis of the impact of cluster number and syllable number on performance. Then, in Section 5.3.3.1.2, I analyse the types of errors made.

5.3.3.1.1. The relationship between cluster number and syllable number

Table 5.10 presents the correct scores according to syllable and cluster number.

Table 5.10. Scores according to cluster number and syllable number

Cluster number	Syllable number		G-SLI	LA1	LA2
0	1	**Mean (SD)**	92.50 (12.08)	82.50 (23.72)	90.00 (12.91)
	2	**Mean (SD)**	70.00 (23.72)	83.75 (14.49)	91.25 (10.29)
	3	**Mean (SD)**	68.75 (27.16)	88.75 (14.97)	93.75 (8.84)
	4	**Mean (SD)**	50.00 (35.36)	67.50 (31.29)	87.50 (17.68)
1	1	**Mean (SD)**	77.50 (14.19)	85.00 (14.19)	86.25 (10.95)
	2	**Mean (SD)**	72.50 (22.48)	81.88 (11.95)	94.38 (6.88)
	3	**Mean (SD)**	53.75 (30.93)	76.25 (10.95)	85.00 (9.86)
	4	**Mean (SD)**	38.75 (34.59)	58.75 (27.67)	81.25 (13.50)
2	1	**Mean (SD)**	62.50 (29.46)	85.00 (17.48)	82.50 (23.72)
	2	**Mean (SD)**	56.25 (22.44)	78.75 (13.24)	93.75 (10.62)
	3	**Mean (SD)**	51.25 (36.54)	71.25 (16.72)	85.00 (14.19)
	4	**Mean (SD)**	37.50 (39.53)	50.00 (31.18)	72.50 (21.89)

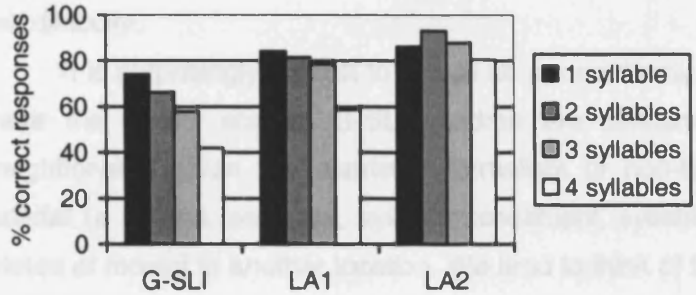
A 3 (Group: G-SLI, LA1, LA2) x 4 (Syllable number: 1, 2, 3, 4) x 3 (Cluster number: 0, 1, 2) ANOVA reveals significant main effects of group, $F(2, 27) = 6.759$, $p = 0.004$, syllable number, $F(3, 25) = 24.260$, $p < 0.001$ and cluster number, $F(2, 26) = 13.588$, $p < 0.001$. Of the interactions, only syllable number x group was significant, $F(6, 52) = 3.480$, $p = 0.004$. Paired sample t-tests show that the effect of cluster number comes from performance being significantly better on non-words with no cluster compared to those with one cluster, $t(29) = 3.455$, $p = 0.002$, better on non-words with one cluster compared to those with two clusters, $t(29) = 2.921$, $p = 0.007$, and better on non-words with no cluster compared to those with two clusters, $t(29) = 3.995$, $p < 0.001$. The pattern of performance with regards to clusters is therefore 0 > 1 > 2.

The syllable number x group interaction is shown in Table 5.11 and Figure 5.6 below.

Table 5.11. % correct scores according to syllable number

Syllable number		G-SLI	LA1	LA2
1	Mean (SD)	74.17 (12.55)	84.17 (14.27)	86.25 (10.22)
2	Mean (SD)	66.25 (19.81)	81.46 (9.44)	93.13 (5.81)
3	Mean (SD)	57.92 (29.75)	78.54 (10.76)	87.92 (7.07)
4	Mean (SD)	42.08 (34.27)	58.75 (25.79)	80.42 (13.04)

Figure 5.6. % correct scores according to syllable number



The syllable number x group interaction was first unpacked using one-way ANOVAs within stimuli of each syllable number. For one-syllable stimuli, the effect of group was not significant. For two-syllable stimuli, the effect of group was significant, F (2, 29) = 10.572, p < 0.001. Post hoc comparisons (Bonferroni-corrected) reveal that the G-SLI group performs significantly worse than the LA1 and LA2 controls, p = 0.045 and p < 0.001 respectively, but that there is no significant difference between the control groups. For three-syllable stimuli there is a significant effect of group, F (2, 29) = 6.724, p = 0.004. The G-SLI group does not perform significantly worse than the LA1 group, p = 0.061, but does score worse than the LA2 group, p = 0.004. Again, there is no signficant difference between the control groups. Finally, for the four-syllable stimuli, the main effect of group is significant, F (2, 29) = 5.515, p = 0.010. The only significant pairwise comparison is between the G-SLI and LA2 group, p = 0.008. Therefore group differences only show up for non-words of two syllables and more.

The interaction was next investigated using a series of t-tests within each group, comparing performance on stimuli with different numbers of syllables. In order to reduce the number of comparisons to be made, I just compared performance on conditions with one versus two syllables, two versus three syllables, and three versus four syllables. For the G-SLI group, the only significant difference is between conditions with three and four syllables, t (9) = 2.863, p = 0.019. For the LA1 group, the only significant difference is likewise between conditions with three and four syllables, t (9) = 3.400, p = 0.008. For the LA2 group, however, the only significant difference is between conditions with two and three syllables, t (9) = 4.038, p = 0.003.

5.3.3.1.2. Error analysis

Although the ANOVA analysis shows that G-SLI children find consonant clusters difficult, it is only by looking at the types of errors they make that we can understand what it is about clusters that they find difficult, and hence begin to work out a phonological explanation for their difficulty.

It is surprisingly difficult to create an error scheme for the TOPhS data, particularly where the lowest scoring G-SLI children are concerned. In a sense, it should be straightforward given the restricted formalism of non-linear phonology – phonological material (a feature, segment, syllabic constituent, syllable or foot) can either be added, deleted or moved to another location. We tend to think of the outputs of child phonology as being structurally simpler than those of adults, and therefore expect material to be deleted. This is essentially what we mean when we say that a complex onset is marked relative to a simplex one: we expect a complex onset to be simplified in child phonology, but not a simplex onset to be made more complex. So we might expect the sorts of errors made by the G-SLI and LA control children to be very straightforward, i.e. cluster simplification. This does indeed occur, to clusters in all positions e.g. ***dəfrɪpələ*** → ***dəfɪpələ*** (LJ), ***prɪlfi*** → ***prɪfi*** (OD), ***dremp*** → ***drem*** (CM). Cluster reduction can occur by deletion of one entire segment, e.g. ***prɪlfi*** → ***pɪlfi*** (SA), or coalescence, the creation of a new segment by combining features of the two segments, e.g. ***prɪlfi*** → ***dɪlfi*** (GD), where */d/* has the stopness of */p/* and the coronality of */r/*. However, not only are clusters reduced, but they are also created, again in each word position, e.g. ***dəfɪpl*** → ***dəfrɪpl*** (GS), ***fəketələ*** → ***fəlketələ*** (LJ), ***ket*** → ***kent*** (QC). Sometimes both cluster reduction and cluster creation occur in the same word, e.g. ***fəkletə*** → ***kəlestə*** (GD). The segmental material from the onset is reattached and there is no place for the initial */f/*. But where does the */s/* come from? And how should we analyse some of the outputs from the lowest-scoring G-children which bear so little resemblance to their targets, e.g. ***səprɪfi*** → ***dəfɪfi*** (GD), where the */d/* possibly comes from the coalescence of */p/* and */r/*, then moved to word-initial position, and then the */s/* has become */f/*, but is this because it retains the labiality from the original */p/* that was in that position, or is it a copy of the onset of the third syllable? It is not uncommon for a non-word to have several errors involving clusters, e.g. deletion and addition, as well as transpositions and coalescence, and it is very difficult to form a neat categorisation. So rather than attempting to create an exhaustive error classification and give statistics for the frequency of occurrence of each, I discuss in the following section just three error types, which I use later in Section 5.5.1 to shed light on the phonological representations of children with G-SLI. It is important to note that these errors do also occur in typically developing children, but less frequently.

The data in (1) and (2) reveal that G-SLI children make errors in the attachment of the second consonant (C_2) of a complex onset, linking it to the wrong onset. Rhymal misattachment, as in (3), is also attested, although it is rarer. Misattachment errors are more common in children with scores in the higher range; those with low scores tend to reduce clusters completely by deleting C_2.

(1) LJ	***fəklɛt* → *fləkɛt***	***fəklɛtə* → *fləkɛtə***	***fəklɛstələ* → *fləlkɛstələ***
(2) CT	***bədrɛp* → *brədrɛt***	***bədrɛpə* → *brədrɛpə***	
(3) GS	***fəklɛstə* → *fəzkɛstə***		

It is also noteworthy that G-SLI children create clusters in non-words that previously lacked them, as in (4).

(4) GS	***dəfɪpl* → *dəfrɪpl***	***pɪfətə* → *prɪfətə***	***dɛpəri* → *dɛmpəri***

(5) and (6) show that on occasion children with G-SLI can realise non-words as real words.

(5) SM	***drɛmpə*** → *jumper*	***klɛsti*** → *crusty*	***fɪpələ*** → *flipper*
(6) GS	***klɛt*** → *collect*	***bədɛp*** → *protect*	***kɛst*** → *kissed*

Note that the target non-word and its real word replacement share properties, such as the majority of segments and metrical structure (on most occasions, but not all: when changes in metrical structure do occur, the output generally has trochaic foot structure). Typically developing children make the occasional lexicalisation but their errors only involve minor changes: ***fəkɛt*** → *forget* and ***kɛst*** → *kissed* are the most common, and ***fɪmpl*** → *simple* and ***pɪlf*** → *pill* are also attested.

The data in (7) and (8) reveal that children's errors are inconsistent. In (7) all three words are matched for metrical and syllabic structure, but in one the word-final consonant is omitted. (8) shows that this optionality occurs even within the same cluster.

(7) CT *səprɪlf* → √ *fəklest* → √ *bədremp* → *bədrem*

(8) SA *demp* → √ *fɪmp* → *fɪm*

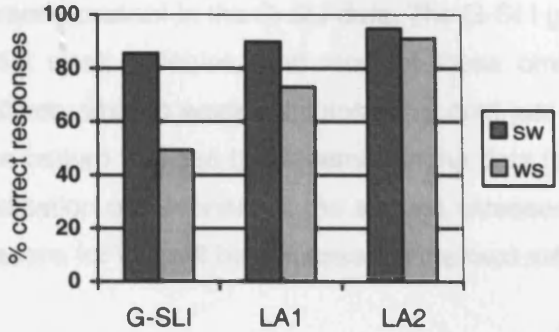
5.3.3.2. Analysis according to metrical complexity

The discussion in the previous section has focused on the impact of syllabic complexity. Addressing the impact of metrical structure is less straightforward because of the confound between metrical structure and syllable number. For example, in the TOPhS stimulus set, all three-syllable non-words have one unfooted syllable and all four-syllable non-words have two unfooted syllables. I have therefore decided to first analyse performance on two-syllable non-words, which have the contrasting metrical patterns weak-strong (e.g. *fəkɛt*) and strong-weak (e.g. *klɛtə*), and then analyse three syllable non-words, with wsw (e.g. *fəkɛtə*) and sww (e.g. *kɛtələ*) patterns. The data are shown in Table 5.12. Although the data are presented together, the results for the two- and three-syllable words are analysed separately.

Table 5.12. Repetition accuracy on two- and three-syllable non-words

Syllable number	Metrical structure		G-SLI	LA 1	LA2
2	sw	Mean (SD)	85.63 (14.44)	90.00 (6.25)	95.00 (4.93)
	ws	Mean (SD)	49.34 (30.25)	73.13 (16.68)	91.25 (9.86)
3	sww	Mean (SD)	63.75 (29.43)	82.50 (11.33)	90.00 (8.44)
	wsw	Mean (SD)	48.13 (33.73)	74.38 (18.27)	84.38 (9.43)

Figure 5.7. Repetition accuracy on two-syllable non-words



A 2 (Group: G-SLI, LA) x 2 (Condition: sw, ws) ANOVA within two-syllable non-words reveals main effects of group, F (1, 27) = 10.280, p < 0.001 and condition, F (1, 27) = 25.869, p < 0.001. In addition, there is a significant group x condition interaction, F (2, 27) = 6.413, p = 0.005, reflecting the fact that the ws condition is particularly difficult for G-SLI children. To unpack the interaction, one-way ANOVAs were carried out in order to investigate each group's performance within each condition. There was no effect of group for the sw condition, but for the ws condition the main effect of group was significant, F (2, 29) = 10.250, p < 0.001. The G-SLI group performs significantly worse than the LA1 group, p = 0.049, and than the LA2 group, p < 0.001, but the difference between the control groups is not significant. To unpack the interaction further, t-tests were carried out comparing performance on the sw and ws conditions within each group. Both the G-SLI and LA1 groups perform worse on the ws condition, t (9) = 3.996, p = 0.03, and t (9) = 3.151, p = 0.012 respectively. The LA2 group does not perform significantly differently on the two conditions. The interaction therefore comes from the G-SLI group having particular difficulty with two-syllable non-words that contain an initial weak syllable.

A 2 (Group: G-SLI, LA) x 2 (Condition: sww, wsw) ANOVA within three-syllable non-words reveals main effects of group, F (2, 27) = 6.999, p = 0.004, and condition, F (1, 27) = 10.834, p = 0.003, but no significant interaction. The main effect of condition arises because performance on the wsw condition is significantly worse than that on sww condition. Post hoc tests for group indicate that the G-SLI group performs significantly worse than the LA1 and LA2 groups, p = 0.044, and p = 0.004 respectively, but that there is no significant difference between the control groups.

Therefore even when syllable number is controlled for, metrical structure has an impact on repetition accuracy in both groups: performance is poor on non-words with a left adjoined syllable, and performance for the G-SLI group particularly so. Why should the presence of a left-adjoined syllable make a difference to repetition accuracy? It is noteworthy that the weak syllable omission that is so characteristic of younger SLI children is rarely present in the G-SLI data. The G-SLI group as a whole omits only 2.92% (14/480) initial weak syllables, and most of these omissions are made by just one child. 6/10 children omit no weak syllables, 2/10 omit just 1, 1/10 omits 3 and 1/10 omits 9. Instead, one pattern that can be discerned in the data is that a left adjoined syllable impacts on the realisation of the onset of the second, stressed syllable, often in complex ways. Possible reasons for this will be discussed in the next section.

5.4. Positional markedness effects for onset clusters: An Optimality Theoretic account

The aim of this section is to begin exploring the impact that metrical structure, specifically the presence of an initial unfooted syllable, has on the realisation of non-words. I focus on an extreme case found in one G-SLI child, GD, who generally realises onset clusters correctly when they are word-initial, but never does so when they occur after an initial weak syllable, i.e. in word-medial position. I show that this pattern of onset realisation is also found to a less striking degree in typically developing children, and in some of the other children in the G-SLI group. The typically developing children whose data are used in this analysis are the same 20 who are described in Section 5.2.2.

In the TOPhS stimulus set, 48 non-words contain an onset cluster. In 24 of these the cluster occurs word-initially, e.g. ***drepə***, and in the other 24 it occurs word-medially, e.g. ***bədrεpə***. I make a distinction between onset clusters which are produced segmentally faithfully and those where a cluster is produced in the target position but which is segmentally unfaithful, e.g. ***dəfrɪp*** → ***dəprɪpf***. The data for GD and the two control groups are presented in Table 5.13.

Table 5.13. Mean (SD) % of onset clusters produced word-initially and word-medially

Onset cluster position		**GD**	**LA1**	**LA2**
Word-initial	**Segmentally correct**	75.00	88.33 (9.78)	95.83 (6.51)
	Segmentally incorrect	4.17	0.83 (2.64)	0.00 (0.00)
	Total	79.17	89.17 (10.05)	95.83 (6.51)
Word-medial	**Segmentally correct**	0.00	75.83 (18.40)	88.33 (11.59)
	Segmentally incorrect	12.50	2.92 (5.22)	0.42 (1.32)
	Total	12.50	78.75 (16.60)	88.75 (10.77)

Table 5.13 shows very clearly that GD is much more likely to retain complex onsets word-initially (e.g. ***drepə***) than word-medially (***bədrepə***). Complex onsets are rarely realised word-medially, and on those occasions the segmental material is invariably altered.

For the control children, a 2 (Group: LA1, LA2) x 2 (Condition: word-initial, word-medial) ANOVA reveals a significant main effect of condition, $F(1,18) = 12.345$, $p = 0.002$, but no significant effect of group or interaction. The main effect of condition results from better performance on word-initial clusters than word-medial clusters. Hence the typically developing children show the same effect of cluster position as GD does, but in a less extreme form.

Now I consider the types of errors made in word-medial position which result in cluster simplification. GD makes three types of errors:-

1) Cluster reduction e.g. ***dəfrɪmp → dəfɪmp, dəfrɪmpələ → dərempfələ.***
2) Vowel epenthesis e.g. ***bədrepə → dərepə, fəkletələ → kəletələ.*** Note that the overall metrical structure of the non-word is unchanged when epenthesis takes place: there are no examples such as ***bədrepə → bədərepə.***
3) Unclassified e.g. ***səprɪfi → dəfɪfi, bədrepəri → dəpɪfəri.*** The first example looks like a case of possible coalescence of */pr/* to */f/* but may just be harmony with the */f/* of the third syllable. The second example looks like deletion of the */r/* in the complex onset and the metathesis of */d/* and */b/*. However, it is not clear that these are the correct explanations, and so these errors will be considered unclassified.

The control children's errors are also classified using this scheme, with the inclusion of no responses in the unclassified category. The means and standard deviations for the three types of errors, expressed as a percentage of total responses, are shown in Table 5.14.

Table 5.14. Mean (SD) % of response types where a word-medial onset cluster is simplified

Error response type	**GD**	**LA1**	**LA2**
Cluster reduction	25.00	12.50 (9.62)	9.17 (8.96)
Vowel epenthesis	33.33	2.08 (4.05)	0.83 (1.76)
Unclassified	29.17	6.67 (8.15)	1.25 (2.01)

For GD, errors are fairly evenly distributed amongst the different error types. For the controls, a 2 (Group: LA1, LA2) x 3 (Error type: cluster reduction, vowel epenthesis, unclassified) revealed a significant effect of error type, $F(2,36) = 14.262$, $p < 0.001$, but no significant effect of group or interaction. Paired t-tests revealed that cluster reduction errors are significantly more common than epenthesis errors and unclassified errors, t (19) = 4.682, $p < 0.001$ and t (19) = 3.676, $p = 0.002$. Unclassified errors are not significantly more frequent than vowel epenthesis errors. Therefore GD shows a different pattern of errors to the control children. While GD produces relatively equal numbers of errors, for the LA1 controls, reduction errors are the most common.

Some of the LA controls show additional evidence of pressure to create an output with the complex onset in word-initial position. On occasion, the cluster is realised instead in the onset of the initial weak syllable, e.g. ***dəfrɪmpələ → drəfɪmpələ***, ***fəkletə → fləkɛtə***, ***bədrepə → brədepə***. On other occasions, the original cluster is retained and a further cluster is created word-initially, e.g. ***fəklestələ → fləklestələ*** and ***bədrep → brədrep***. There are also occasions when the initial weak syllable is deleted so that the complex onset is now word-initial: ***fəklestə → klestə***. GD, however, does not make these types of errors. These data show that even though word-position markedness effects are not as strong in the LA children as they are for GD, they are still in evidence.

Why are GD and typically developing children more likely to realise a complex onset correctly when it is word-initial? Here I present a positional markedness account of the data within an OT framework. The aspects of the data that the account needs to capture are:-

- Both GD and the LA children generally realise word-initial onset clusters correctly.
- GD simplifies word-medial onset clusters by either reduction or vowel epenthesis.
- LA children also simplify word-medial onsets, but also realise a large proportion correctly.

The account I present here makes use of two markedness constraints that refer to complex onsets: a general markedness constraint, *COMPLEXONSET, and a specific markedness constraint, LICENSECOMPLEXONSET. I define these constraints as follows:-

***COMPLEXONSET** – onset clusters are not licensed.

LICENSECOMPLEXONSET – onset clusters are only licensed word-initially.

Also needed is the faithfulness constraint **MAX-C**, which requires the output to be faithful to consonants in the input.

Let's start with the strong positional markedness effects found in GD's grammar. A ranking of LICENSECOMPLEXONSET >> MAX-C >> *COMPLEXONSET achieves this pattern of onset behaviour, as shown in Tableaus 5.1 and 5.2.

Tableau 5.1. GD's grammar

	/frɪmp/	LicenseComplexOnset	Max-C	*ComplexOnset
a. ☞	*frɪmp*			*!
b.	*fɪmp*		*!	

Tableau 5.2. GD's grammar – absolute positional markedness effects

	/dəfrɪmp/	LicenseComplexOnset	Max-C	*ComplexOnset
a.	*dəfrɪmp*	*!		*
b. ☞	*dəfɪmp*		*!	
c. ☞	*fərɪmp*		*!	

For typically developing children, however, positional markedness effects are present, but word-medial cluster simplification occurs alongside correct production. In order to account for this variability, LicenseComplexOnset and Max-C must be equally ranked. This ranking is shown in Tableau 5.3.

Tableau 5.3. LA grammar – variable positional markedness effects

	/dəfrɪmp/	LicenseComplexOnset	Max-C	*ComplexOnset
a. ☞	*dəfrɪmp*	*!		*
b. ☞	*dəfɪmp*		*!	
c. ☞	*fərɪmp*		*!	

In adult grammar there are presumably no positional markedness effects on onset clusters, given that words such as *afraid*, *detract* and *applaud* exist. Max-C must therefore be ranked above the two markedness constraints, as shown in Tableau 5.4.

Tableau 5.4. Adult grammar – no positional markedness effects

	/dəfrɪmp/	Max-C	LicenseComplexOnset	*ComplexOnset
a. ☞	*dəfrɪmp*		*!	*
b.	*dəfɪmp*	*!		
c.	*fərɪmp*	*!		

I next look at the other children in the G-SLI group to determine whether they show positional markedness effects too. Table 5.15 presents the relevant data for each of the other nine G-SLI children, and those that show positional markedness effects are highlighted in red.

Table 5.15. Performance on initial and medial onset clusters in the G-SLI group.

Child	TOPhS score (/96)	Initial clusters segmentally accurate	Initial clusters segmentally inaccurate	Medial clusters segmentally accurate	Medial clusters segmentally inaccurate
LJ	71	24	0	15	1
CM	60	22	0	14	0
SA	77	21	0	15	0
SM	34	14	2	7	4
CT	**81**	**21**	**2**	**22**	**0**
GS	**38**	**18**	**1**	**18**	**3**
OD	**32**	**13**	**0**	**10**	**0**
QC	**79**	**21**	**0**	**22**	**0**
BD	**84**	**21**	**0**	**21**	**0**

Only 4/9 children show clear positional markedness effects. This split is not a function of overall performance on the TOPhS: the three lowest scorers are SM, GS and OD, but of those, only SM shows positional markedness effects.

Interestingly, these positional markedness effects manifest themselves in different ways. LJ realises the cluster word-initially instead, e.g. ***bədrɛpə*** → ***drədrɛpə***, ***fəkletə*** → ***fləketə***. SA reduces clusters e.g. ***dəfrɪmpl*** → ***dəfɪmpl***, ***səprɪʃi*** → ***səpɪʃi***, and CM makes a mixture of both types of error, e.g. ***bədrɛpə*** → ***drəpɛpə*** and ***dəfrɪmpl*** → ***dəfɪmpl***. SM generally reduces clusters, e.g. ***səprɪʃi*** → ***səpɪsi***. SM also makes substitutions of a cluster by a single segment that does not appear to be a product of coalescence, e.g. ***dəfrɪpl*** → ***kəpɪpl***, ***bədrɛmp*** → ***bəfɛmp***, and which are in that respect reminiscent of GD's unclassified errors. It is not clear why this variation in error types occurs. Different errors may represent different strategies for dealing with the same underlying problem. Alternatively, there may just be more variation in individual phonological grammars than is usually recognised (see Tzakosta, in prep.).

Finally, I stress that positional markedness effects on onset clusters have not been previously noted in the literature. The data from the LA control children suggest that such effects are a feature of typical development. I propose that children's acquisition of onset clusters proceeds in stages, with LICENSECOMPLEXONSET originally ranked above MAX-C, which is itself ranked above *COMPLEXONSET. LICENSECOMPLEXONSET is then gradually demoted, thereby accounting for the disappearance of these effects over time. Of course, this sequence needs to be confirmed through the collection of data from typically developing children who are younger than those who participated here. Even then, outputs may not look like GD's, because young English-speaking children have high rates of initial weak syllable omission: whereas GD produces outputs such as *dəfɪmp* for *dəfrimp*, typically developing children might produce *fɪmp* or *frɪmp*. This proposed interaction with initial weak syllable omission might explain why such patterns haven't been picked up before. It is possible that the study of disordered phonology, where clear patterns such as GD's may be evident, can alert researchers to phenomena which may also be found in typical development.

5.5. Discussion

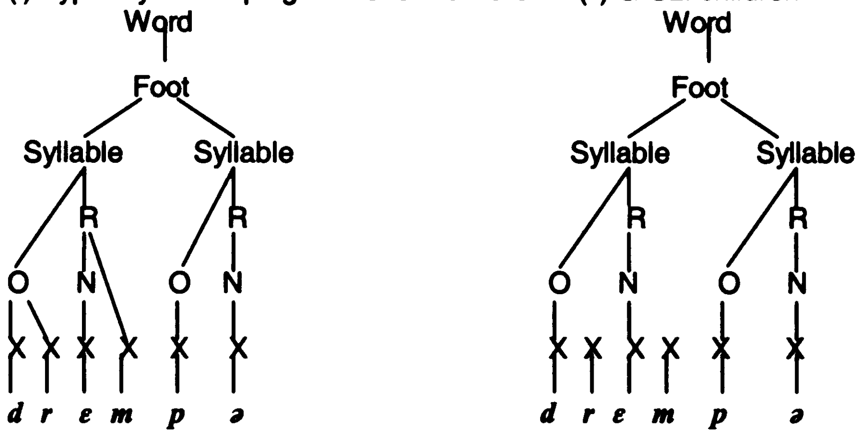
5.5.1. A model of phonological representations in G-SLI

In Section 5.3.3.1.2. I discussed some of the errors that G-SLI children make when repeating non-words with clusters. In addition to simplifying clusters, G-SLI children also create them. The error I term 'consonant misattachment' (e.g. *fəklet* → *fləket*) is one I interpret as revealing a difficulty in joining up the additional consonant to the prosodic hierarchy. The child knows that this consonant has to go somewhere in the word, but cannot remember where. Similarly, G-SLI children create clusters in non-words that previously lacked them, as though they were carrying over C_2s and rhymal consonants from previous non-words. Such errors, which result in an increase rather than a decrease in complexity, have also been noted by Weismer and Hesketh (1996) when teaching SLI children novel words. They concluded, as I do, that syllabic errors reflect more than just a tendency to reduce the form to one that is easier to articulate (c.f. Bishop *et al.*, 1996). Another error type, lexicalisation (whereby children repeat non-words as real words) has been noted in previous studies of non-word repetition (e.g. Dollaghan *et al.*, 1995; Stackhouse, 1993; Weismer & Hesketh, 1996). I suggest that real words are well-rehearsed sound sequences, and therefore pose less of a load on the memory (Marshall *et al.*, 2002). Finally, the optionality in children's production of clusters shows that it is not the case that children *never* produce a particular structure. This optionality mirrors that found for syntax (e.g. van der Lely, 1998) and inflection (van der Lely & Ullman, 2001). Complex grammatical structures are not unavailable, merely only optionally available

I propose a model of syllabic representation in G-SLI whereby children have branching onset and branching rhyme structure only optionally available to them. Figure 5.8(i) shows that for typically developing children the additional /r/ and /m/ of *drɛmpə* can be joined to prosodic hierarchy. This structure is available to G-SLI children some of the time, and when it is the non-word is realised correctly. Figure 5.8(ii) shows that the branching structure is not always available to G-SLI children. When it is not available, the additional consonants /r/ and /m/ cannot be joined to the prosodic hierarchy, meaning that they cannot be licensed, and therefore cannot be realised.

Figure 5.8. Representations of syllabic complexity in:

(i) Typically developing and G-SLI children (ii) G-SLI children



Depending on whether branching syllabic structure is available, G-SLI children will sometimes realise a syllabically complex word correctly and sometimes not. In contrast to previous work on SLI (Bishop *et al*, 1996; Gathercole & Baddeley, 1990), I have shown that for G-SLI children, cluster errors are not just in the direction of simplification. The finding that clusters are created as well as reduced reveals that the difficulty is one of representation, rather than one of poor articulation. Intriguing is the discovery that word position affects cluster accuracy (Section 5.4). Any model of syllabic representations in both typically developing and G-SLI children will have to take this phenomenon into account. Metrical structure seems to be properly represented in G-SLI children, in the sense that unfooted syllables are not omitted. However, given that the TOPhS stimuli have a confound between metrical structure and word position (medial onset clusters invariably follow an initial weak syllable), I am still pondering the status of metrical representations in G-SLI and how this might affect cluster accuracy. I have no answers at this stage.

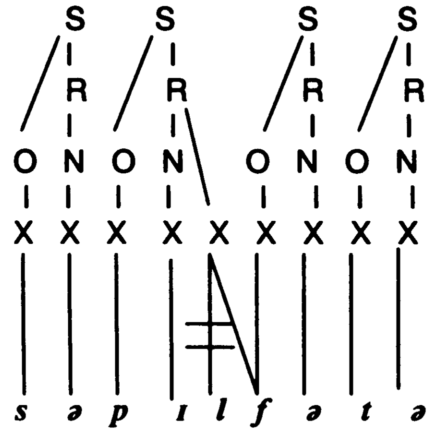
However, we can now explain why G-SLI children perform poorly on longer words in non-word repetition tests. It is *not* that they have limited capacity in their phonological store compared to other children. Rather, words cannot always be assigned a full structural representation, leading to errors in repetition. The causal arrow is reversed.

Given that some G-SLI children *do* have language-age appropriate scores, does this mean that these children have normal syllabic representations of clusters, albeit with the same degree of optionality that is demonstrated by typically developing children? On my model, this would be the case. However, the issue is an empirical one, and it is not clear how I can resolve it with the data here. Certainly it is possible that G-SLI children represent clusters differently to typically developing children (see Marshall *et al.*, 2003). Indeed, there are precedents for claiming that two groups can show similar overt behaviour but possess different cognitive mechanisms, as in face-processing and lexical learning in Williams Syndrome and typically developing individuals (Deruelle, Mancini, Livet, Casse-Perrot & de Schonen, 1999; Stevens & Karmiloff-Smith, 1997).

One way of tackling this particular issue with the data I have here would be to compare the types of errors made by the language-age appropriate G-SLI scorers and their controls. It is possible that there are both quantitative and qualitative differences in the errors made by the language-age and below language-age scorers, but this can not be investigated without an understanding of how to accurately characterise errors. Unfortunately we are far from being able to create an error-analysis scheme for the TOPhS, for reasons discussed in Section 5.3.3.1.2.

What about BD, who appears to have chronological age-appropriate phonological representations? Does he really have access to intact syllabic representations? Two of BD's errors involve compensatory lengthening, whereby a segment spreads to fill the slot left vacant when a previous segment is deleted. ***səpɪlfətə*** becomes ***səpɪffətə*** and ***səprɪlfətə*** becomes ***səpɪffətə***, and the first of these is illustrated in Figure 5.9. The */l/* has been deleted. Due to the independence of melody and prosody, a consonant slot is left, into which neighbouring segmental material, in this case the /f/ of the following onset, can spread. The result is a geminate, and English does not have geminates inside monomorphemic words, although they can be formed by prefixation (e.g. *unnatural*) and compounding (*night time*). However, the important point is that although the output is unusual, the process of compensatory lengthening shows that syllable structure is intact. If BD had no facility for representing branching rhymes, we would not expect */f/* to spread after the deletion of the */l/*. Importantly, there are no gemination errors when there is no spare X slot to fill, e.g. of the type ***səpɪfətə*** → ***səpɪffətə***, so when gemination does occur it cannot satisfactorily be explained as an articulation error. Therefore I conclude that BD does have intact prosodic structure.

Figure 5.9. Gemination in BD's data



5.5.2. What individual variation means for theories of SLI

One theoretically important finding to come out of this study is that, while phonological skills are impaired in the majority of the G-SLI population, not every G-SLI child has poor phonology. BD has age-appropriate phonological abilities as measured by the TOPhS, despite his severe deficits in syntax and vocabulary. Gallon *et al.* (2004) have investigated the performance of a group of thirteen G-SLI children on the TOPhS. They found a wide range in scores, from a low of 31% to a high of 98%. The existence in Gallon's group of one child with a chronological age-appropriate score (the child who scores at 98%) shows that BD's unimpaired phonology is not unique. Together these data show that a phonological impairment is not a necessary part of the linguistic profile of G-SLI. The syntactic impairments characteristic of this group can occur in the absence of phonological impairments. Similarly, and in a much larger sample of younger SLI children, Conti-Ramsden, Botting and Faragher (2001) found that only 78% of the group had problems with the CNRep relative to their chronological age.

On the other hand, I reported in Section 5.3.2.3 on WD, who has normal language abilities but performs poorly on the TOPhS. Poor phonology is neither necessary nor sufficient to cause G-SLI. Furthermore, phonology is correlated neither with age, nor grammatical ability nor vocabulary ability (see Section 5.3.1). It is therefore difficult to see how phonological difficulties can be argued to underlie G-SLI. Similar points have been made by Rosen (2003) and van der Lely, Rosen and Adlard (in press). Difficulties with phonological complexity appear to be neither necessary nor sufficient for the grammatical problems seen in the G-SLI subgroup.

In the remaining G-SLI children we find two patterns of performance – language age-appropriate and below language age-appropriate. This is the same conclusion as that reached by Ebbels in a study of 15 children with severe mixed SLI (Ebbels, 2003). She finds a bimodal distribution of TOPhS scores, with the children divided between a low scoring group (33.33% - 48.96%) and a high scoring group (70.83% - 92.71%). The low scoring group score significantly worse than their grammar and vocabulary-matched controls, whereas the high scoring group score language age-appropriately.

An important issue in psycholinguistics is what children with linguistic behaviour far from the group mean tell us, together with the related issue of whether we should ignore those outliers. As I see it, there are two rather different questions here: (1) What is the most common behaviour of the group? (2) What is the range of possible behaviours of the group? The work presented in this chapter, in line with that of Ebbels *et al.* (2003) and Gallon *et al.* (2004) shows that many, if not most, children with G-SLI do have difficulties with phonology. This is of clinical importance – it indicates that many of these children will need phonological therapy. It is also of theoretical importance, because it suggests that for these children the existence of a causal link between poor phonology and poor grammatical skills is worth pursuing.

On the other hand, the total range of variation within the group tells us what is and what is not possible, and therefore what needs to be accounted for by our theories. If a theory says that phonological difficulties are a necessary precursor to syntactic difficulties, and yet we identify a child who has syntactic difficulties but no phonological difficulties (at least, as measured by this particular phonological test), then that theory is seriously undermined and needs to be refined.

Establishing what is responsible for the heterogeneity in language profiles in the SLI population is not trivial. Language is a complex system. Some processes are likely to be domain-specific, others more general processes shared with other aspects of cognition, and it is likely that within the population as a whole more than one of these processes can be impaired. Furthermore, how can we be sure that this heterogeneity comes from different underlying deficits rather than from different patterns of compensation, particularly in a group such as this that has had years of intensive speech and language therapy?

The question of whether the deficits in SLI form a single linguistic core problem or several separate problems is perhaps 'the most interesting for linguistics and language acquisition' (de Villiers, 2003:431). Work over the last ten years by van der Lely and colleagues provides evidence that SLI is a heterogeneous disorder, but that a subgroup with relatively homogeneous grammatical weaknesses, the G-SLI subgroup, can be identified. The results reported in this chapter suggest that G-SLI children are not homogeneous in relation to their phonological abilities. A hetereogeneous picture also emerges when investigating the impact of phonology on morphology in future chapters. I have therefore chosen in this thesis to report behaviours characteristic of the group as a whole, but also to discuss individuals who show patterns different to the group mean, because I believe this is the only way to obtain a full characterisation of G-SLI grammar(s). The finding that different aspects of language can be differentially impaired in different children is evidence that the language faculty has a modular architecture. Providing a precise description and explanation of how components of language break down in SLI is critical to the development of the Deficit in Computational Grammatical Complexity Hypothesis that has been the focus of van der Lely and colleagues' work over the past three years, and that I develop in subsequent chapters.

PART 3.

INVESTIGATING THE IMPACT OF A PHONOLOGICAL DEFICIT ON PAST TENSE INFLECTION

Chapter 6. The impact of verb-end complexity on the judgement of past tense forms

6.1. Introduction

6.1.1. Chapter outline

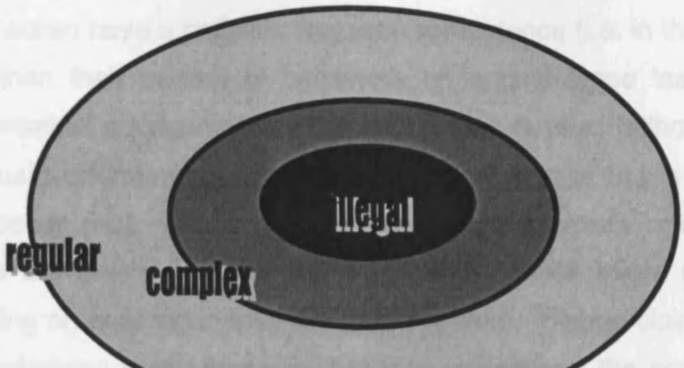
In Chapter 5 I showed that the majority of children with G-SLI have difficulty representing prosodically complex structures in non-word repetition tasks. Children who have difficulty representing complexity at the word-end are predicted to omit suffixes when the addition of a suffix would create a cluster, and I test this prediction later in the experiments reported in Chapter 7. In *this* chapter I investigate how a difficulty in representing prosodic complexity impacts on grammaticality judgements of past tense inflection. The experiment reported here explores the relationship between verb-end complexity in both regular and irregular verbs, and successful judgements of their past tense forms. It uses only regular verbs that are phonotactically legal, because I have already shown, in Chapters 3 and 4, that phonotactics affect past tense inflection independently of prosodic complexity.

In Section 6.1.2 I discuss the relationship between verb-end phonotactics and verb-end complexity. In Section 6.1.3 I discuss the use of grammaticality judgements in psycholinguistic research. In Section 6.2 I present the method, and in Section 6.3 the results. I summarise and discuss the results in Section 6.4.1, and consider the impact of prosodic complexity on past tense judgements in more detail in Section 6.4.2.

6.1.2. Verb-end complexity

Figure 6.1 illustrates the relationship between regular morphology and two dimensions of verb-end phonology – phonotactic legality and prosodic complexity. As the figure shows, these dimensions are not orthogonal. Within the set of verbs that take regular past tense morphology, some, represented by the light grey shading, have inflected endings that are phonotactically legal (i.e. the same endings are found in monomorphemic words) and prosodically simplex (i.e. they do not contain a cluster). Examples of such verbs are *played* and *sewed*. Some regular verbs have prosodically complex endings, and they further divide into two groups depending on the phonotactic legality of that ending. Some have a cluster that is legal (dark grey shading), and examples include *yelled* and *tossed*. The remainder have a cluster that is illegal (black shading), with examples including *hugged* and *danced*.

Figure 6.1. Prosodic complexity and phonotactic legality in regular past tense verbs



This investigation considers not only the prosodic complexity of the inflected verb end, i.e. whether or not the verb ends in a cluster, but also the length of the preceding nucleus. For an inflected verb whose only verb-end consonant is the suffix, the preceding nucleus must be long because of a constraint in English on the minimal size of monosyllabic words – a verb (in common with other content words) has to be minimally a foot, which means that if it ends in a vowel, then that that vowel has to be long (see Section 1.2.3.1). This constraint on the length of the nucleus is not present if the verb stem ends in a consonant. In that case, the nucleus can be either short or long.

6.1.3. Grammaticality judgement tasks

The task is a forced-choice grammaticality judgement task. Judgement tasks are widely used in psycholinguistics to inform us about children's grammatical knowledge (McDaniel & Smith Cairns, 1996). They have been used to test knowledge of such syntactic aspects as binding theory, WH-movement, relative clause constructions and subject-auxiliary inversion (see references in McDaniel & Smith Cairns, 1996), as well as morphology (e.g. Montgomery & Leonard, 1998; Rice, Wexler & Redmond, 1999). Two types of task are used – open-ended and forced choice. In an open-ended task, the experimenter presents a sentence, e.g. **Whose did you read book*, and the child says whether or not it sounds right. One disadvantage of open-ended tasks is that children have a bias to say 'yes' in their responses, which means they may accept an ungrammatical answer as correct even though they know that it is ungrammatical. This yes bias may reflect a social bias towards acceptance (McDaniel & Smith Cairns, 1996). In a forced-choice task the child has to choose which of two sentences is the right one, e.g. *Whose book did you read*/ **Whose* *did you read book*. A disadvantage of forced choice tasks is that children have to choose just one of the two sentences, when in fact both might be acceptable to them.

If SLI children have a deficit in linguistic competence (i.e. in their actual knowledge of language), then their pattern of behaviour on a past tense task should be similar regardless of whether a judgement or production task is used (although the type of task may affect actual performance levels). On the other hand, if SLI is caused by a limited processing capacity (e.g. Bishop, 1994), then these children's underlying grammatical representations are intact, and the differing task demands might give rise to different results depending on how much they stress the system. Bishop claims that 'SLI children do have the underlying competence in that they understand the grammatical function of morphological markers, but fail to apply their knowledge consistently because of limitations on their processing capacity.' (Bishop, 1994:508). This suggests that children with SLI should not be impaired for tense when tested on judgement tasks.

Few judgement tasks have been conducted to test SLI children's morphosyntactic abilities, in contrast to the large number of elicitation tasks. Those that have been carried out with past tense stimuli have indicated that SLI children have difficulties with judgement, although it is not clear that they perform worse than their language-matched peers. In a longitudinal study of SLI children aged between 6;00 and 8;00, Rice *et al.* (1999) found that they only performed worse than language-matched controls at some time periods, and that they otherwise performed at the same levels. Montgomery and Leonard (1998) found that SLI children aged 8;06 performed significantly worse than age-matched controls but not MLU-matched controls. Van der Lely and Ullman (1996) showed that children with G-SLI judge stem forms like **walk* and overregularisations such as **falled* to be acceptable in past tense contexts. Taken together, these results indicate that SLI children's difficulties with past tense formation are not confined to production, *contra* Bishop (1994).

As far as I am aware, the task reported in this chapter is the first to consider the impact of verb-end complexity on past tense judgements. A representation of prosodic complexity is essential when judging whether or not a regular verb is inflected. The tense change in *roll* → *rolled* is signalled by the formation of a cluster at the verb end. Children who have difficulty representing clusters are predicted to have more difficulty judging *rolled* as the past tense form when compared to children who have no difficulty with clusters. However, it is not predicted that such difficulties in representation will lead to less accurate judgement of irregular past tense forms. Although the tense change in *sell* → *sold* does involve the formation of a cluster, the tense change is also signalled by the change in vowel quality. The child should be able to judge the tense of the verb solely by using the vowel cue, regardless of the accuracy of the verb-end representation.

6.2. Method

6.2.1. Verb stimuli

Six conditions are used in this task, comprising regular and irregular verbs of different degrees of verb-end complexity. Their characteristics are shown in Table 6.1.

Table 6.1. Phonological and morphological characteristics of the stimuli

Condition	Morphology	Phonological characteristics	Examples
VV-D	Regular	Long vowel, no cluster	*played*
VC-D	Regular	Short vowel, 2-consonant cluster	*yelled*
VVC-D	Regular	Long vowel, 2-consonant cluster	*frowned*
VVD	Irregular	Long vowel, no cluster	*made*
VCD	Irregular	Short vowel, 2-consonant cluster	*held*
VVCD	Irregular	Long vowel, 2-consonant cluster	*found*

Regular and irregular stimuli are chosen to match in terms of the segmental content of their past tense ending. Note that the irregular verbs signal tense in different ways: by a consonant change (*make* → *made*), a vowel change (*hold* → *held*) or a vowel change plus addition of */d/* (*tell* → *told*). Hence although their past tense forms have the same complexity as the set of regular inflected forms, an increase in complexity does not play a necessary part in signalling tense in any of these irregular verbs. Eight regular and eight irregular verbs are used in the task, and these are presented in Table 6.2.

Table 6.2. Verb stimuli

Condition	Regular	Frequency*	Irregular	Frequency
VV-D/ VVD	*played*	4.190	*made*	6.146
VV-D/ VVD	*purred*	0.000	*heard*	4.868
VV-D/ VVD	*sewed*	0.000	*rode*	3.714
VC-D/ VCD	*yelled*	3.091	*held*	4.836
VC-D/ VCD	*tossed*	3.135	*lost*	3.912
VC-D/ VCD	*stepped*	3.526	*slept*	2.944
VVC-D/ VVCD	*rolled*	3.555	*told*	5.659
VVC-D/ VVCD	*frowned*	2.079	*found*	5.595

*Frequency values from Francis and Kucera (1982); calculated as ln (raw frequency + 1).

The number of verbs is limited because of the need to match regular and irregular verb ends for exact segmental content, and the constraints on constructing sentences that are syntactically and pragmatically appropriate. The rationale for matching for the segmental content of the verb end is that if complexity is found to affect performance on one set of verbs and not another, we can be sure that this is indeed an effect of complexity rather than of differing segmental content. Unfortunately, with such a small sample of verbs to choose from, it is not possible to control for onset complexity: the set of regulars contains three verbs with an onset cluster and the set of irregulars contains just one. However, in the regular set these onset clusters are distributed evenly amongst the three conditions, and so if an effect of verb-end complexity is found for regular verbs, this effect will be independent of onset complexity.

Previous studies show that frequency impacts on irregular inflection in both typically developing and SLI children, and on regular inflection in SLI (Ullman & Gopnik, 1999; van der Lely & Ullman, 2001). It was not possible to balance the regulars and irregulars for frequency, as irregulars are on the whole considerably more frequent than regulars (see Table 6.2). Instead, frequency was entered into the analysis as a continuous variable.

6.2.2. Procedure

Pairs of sentences are presented to the child in a forced-choice judgement task. One of each pair contains the past tense form of a verb, and the other contains the uninflected stem form, e.g. 'Yesterday I *played* at home'/ '*Yesterday I *play* at home'. The only difference between the two pairs of sentences is the presence/absence of inflection on the verb.

The sentences are presented on a laptop computer as part of a game that has two cats as its main characters. One cat says one of the sentences in the pair, and the other cat says the other sentence. The child has to choose which cat has said something that 'sounds right'. He presses one of two buttons on the computer keyboard to register his choice, and the chosen cat gets a 'reward'. The aim of these rewards is to reduce pressure on the child – it is the cat that is right or wrong, not the child.

The experiment has three parts – an introduction, a practice phase and an experimental phase. The appearance of each image on the computer screen is controlled by experimenter, so that the experiment proceeds at the child's pace. The orange cat is always on the left hand side of the screen and the grey cat is always on the right hand side. Correct and incorrect sentences are allocated to the two cats in random order. The introduction is given by the experimenter as follows:-

<u>On screen</u> – *orange cat sitting and grey cat sitting*

'This is a game with two cats – an orange cat and a grey cat. The cats are going to tell you what they did yesterday. When the orange cat is talking he will stand up and he'll tell you something that he did yesterday.

<u>On screen</u> – *orange cat standing and grey cat sitting*

Listen to what he says. Then the grey cat will stand up and tell you something that *he* did yesterday.

<u>On screen</u> – *orange cat sitting and grey cat standing*

Now you need to think carefully. If you think that what the orange cat said sounds right, then you need to press the orange button (experimenter indicates to the child to press the orange button).

<u>On screen</u> – *orange cat eating fish and grey cat sitting with a sad face*

The orange cat likes fish, so if he's right he gets a fish to eat. Or, if you think what the grey cat said sounds right then you need to press the grey button (experimenter indicates to the child to press the grey button).

<u>On screen</u> – *orange cat sitting with a sad face and grey cat drinking milk*

The grey cat likes milk, so if he's right he gets a bowl of milk to drink.

<u>On screen</u> – *orange cat sitting and grey cat sitting, i.e. back to first picture.*

Before the game starts we'll have a few practice turns so that you can see what you need to do.'

There are six pairs of sentences at the practice phase, grouped into two sets of three. Verbs in four of the pairs are irregular verbs with very salient differences in the stem and past tense forms. None of these verb-endings appear in the experimental phase. The other two pairs are filler sentences which require judgement of verb argument structure.

There are seventy two pairs of experimental sentences, grouped into six sets of twelve. After each set comes a short motivating cartoon. All sets are obligatory. Each verb appears in three pairs of sentences. The rationale for having the child judge each verb three times is that it allows optionality to be investigated. We already know that G-SLI children produce differing forms of the same verb in the same context. For example, in a narrative task children produced forms such as *fell/ *felled, *fall/ fell, *come/ came, *come/ *comed, *look/ looked* in past tense contexts just a few sentences apart (van der Lely, unpublished data). Including each verb three times in the judgement task allows us to determine whether this optionality holds for judgement too. Recall that one of the disadvantages of forced choice judgement tasks is that both forms may be grammatical for the child. Using multiple instances of a particular verb allows the experimenter to investigate whether this is indeed the case.

Stimulus sentences are constructed so that the verb is always followed by a word that begins with a vowel, e.g. 'Yesterday I frowne<u>d</u> <u>a</u>ll day'. Guy (1991) has explored phonetic factors that can lead to the deletion of word-final consonants, particularly */t/* and */d/*, and shown that they are more likely to be deleted when the next word begins with a consonant. This is because stops have only weak, or even no, internal acoustic cues during the closure phase. An important cue to stops is their release burst, whose audibility depends on the nature of the following segment, and this release burst is at its most audible before a vowel (see also the discussion in Cote, 2002).

Forty eight sentences were generated (plus 24 fillers requiring the judgement of argument structure). The order of these was randomised, and then checked to ensure that similar-sounding verbs were not adjacent to one another. The presentation order was identical for each child, and the stimuli are listed in Appendix B.1.

6.2.3. Participants

16 G-SLI children aged 8;11 to 16;07, and 32 typically developing children aged 4;02 to 10;01, participated in the study. Individual matches were chosen for the G-SLI children, such that each G-SLI child had one receptive grammar control (matched on identical raw score for the TROG) and one receptive vocabulary control (matched on raw score ± 3 for

the BPVS). Details of the language matches are given in Table 6.3, with the relevant scores highlighted in red.

Rather than divide up the control children by age, as I did in Chapters 4 and 5, I kept them in language-matched groups, one matched for grammar and one matched for vocabulary. The reason for this is that one of the most pertinent aspects of research into past tense inflection is whether SLI children have a deficit relative to their language matches. This has been shown for production (e.g. Rice & Wexler, 1996), but the picture is less clear for judgement, with group differences being found at some ages but not at others (Rice *et al.*, 1999). In this task then there is a need for closer language matches than was the case in Chapters 4 and 5.

Table 6.3. Participant details

Measure		**G-SLI** N = 16	**Grammar controls** N = 16	**Vocabulary controls** N = 16
Age	**Mean**	13;01	7;04	7;10
	Range	8;11 – 16;07	4;02 – 10;00	4;01 – 10;01
TROG	**Raw, mean**	14.06	14.06	14.80
	Raw, range	6 – 18	6 – 18	8 – 19
	z-score, mean	-1.35	-0.14	0.01
BPVS	**Raw, mean**	83.50	77.30	82.68
	Raw, range	47 – 107	53 – 97	44 – 109
	z-score, mean	-1.89	0.37	0.23

6.2.4. Predictions

Predictor variables and their expected impact on performance are as follows:-

- **Group** – Previous studies report that SLI children perform less accurately than language-matched controls on tasks of past tense inflection (Rice & Wexler, 1996; van der Lely & Ullman, 2001). The grammar and vocabulary control groups are therefore predicted to perform better than the G-SLI group.
- **Verb type (i.e. regular/irregular)** – For G-SLI children, van der Lely and Ullman report no difference in performance on regular and irregular verbs, but they find an advantage for regulars in typically developing children (van der Lely & Ullman, 2001). It is not clear that this regularity advantage will be evident for the control groups in a judgement task, because the sound changes that signal the stem/past distinction are less salient

in regulars than they are in irregulars. I therefore make no prediction as regards the impact of verb type on performance.

- **Verb-end complexity** – Given that prosodic complexity impacts on the phonology of children with G-SLI (see Chapter 5), performance is predicted to be affected by verb end complexity as follows:-
  - For the G-SLI group, increased verb-end complexity will lead to a decrease in accuracy on pairs of regular verbs but not on irregulars. This is because (i) a representation of verb-end complexity is essential for the judgement of inflection in regular verbs but not irregulars and (ii) the G-SLI group has difficulty with representing complex prosodic structures.
  - For the control groups, increased verb-end complexity will not affect performance on either regulars or irregulars, because these children do not have difficulties in representing prosodic complexity.
- **Frequency** – Performance is predicted to improve as frequency increases.

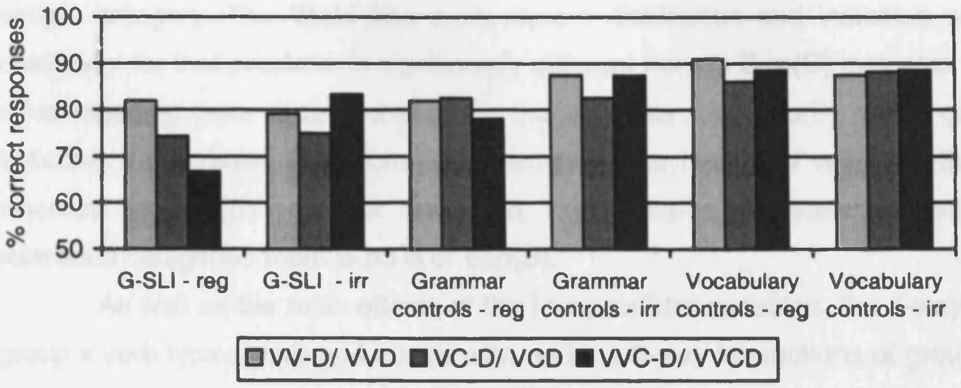
6.3. Results

Scores were automatically coded and recorded using the E-prime experimental programme. Correct scores are coded as 1, incorrect scores as 0. The percentage of correct responses for each participant group and verb type is shown in Table 6.4 and Figure 6.2.

Table 6.4. % correct responses

Condition			G-SLI	Grammar controls	Vocabulary controls
Regular	**VV-D**	**Mean (SD)**	81.94 (38.60)	81.94 (38.60)	91.00 (28.76)
	VC-D	**Mean (SD)**	74.31 (43.85)	82.64 (38.01)	86.11 (34.70)
	VVC-D	**Mean (SD)**	66.67 (47.39)	78.12 (41.56)	88.54 (32.02)
Irregular	**VVD**	**Mean (SD)**	79.17 (40.75)	87.50 (33.19)	87.50 (33.19)
	VCD	**Mean (SD)**	75.00 (43.45)	82.64 (38.01)	88.19 (32.38)
	VVCD	**Mean (SD)**	83.33 (37.46)	87.50 (33.25)	88.54 (32.02)

Figure 6.2. % correct responses



The data were analysed using binary logistic regression. Because this statistical technique might be unfamiliar to the reader, I will explain it in some detail here. Logistic regression takes into account chance performance at 50% (i.e. participants had a choice of two answers), which ANOVA and multiple regression do not allow. The method is basically multiple regression with an outcome variable that is a categorical dichotomy (Field, 2000). A 3 (Group: G-SLI, grammar control, vocabulary control) x 2 (Verb type: regular, irregular) x 3 (Verb-end complexity: VV-D, VC-D, VVC-D) design was used, and frequency was entered as a continuous variable. In a logistic regression, one value for each predictor variable is chosen as the reference category against which other values of that variable are compared. I chose the G-SLI group, regular verbs and VVC-D verb-end as the reference categories.

There are two methods for entering predictor variables into the model. The **forced entry method** is the default method, whereby all variables are placed into the regression model in one block and parameter estimates calculated for each block. In **stepwise methods** the initial model includes only a constant, and then single predictors are added into the model based on their significance (forward stepwise) Alternatively, the model initially includes all predictors, which are then removed one by one to test which ones can be removed without having a substantial effect on the fit (backwards stepwise). Because stepwise techniques are influenced by random variation in the data and seldom give replicable results if the model is retested within the same sample (Field, 2000), I use the entry method in this analysis. Although stepwise models are defensible where no previous research exists on which to base hypotheses for testing, it is clear from Section 2.4 that such research does exist.

I report three statistics – B, the Wald and Exp(B). **B** is the same as the B value in linear regression: it is the value needed to establish the probability that a case falls into a certain category. The **Wald** has a chi-square distribution and indicates whether the B coefficient for that predictor is significantly different from 0. **Exp(B)** indicates the change in odds resulting from a unit change in the predictor. An Exp(B) value greater than 1 indicates an increase in performance relative to the reference value for that parameter, whereas an Exp(B) value of less than 1 indicates a decrease in performance. For reference categories there is no B or Exp(B).

As well as the main effects of the four predictor variables, the 3-way interaction of group x verb type x verb-end complexity, and the 2-way interactions of group x verb type, group x verb-end complexity and verb type x verb-end complexity were investigated. For clarity, the main effects are presented in Table 6.5 and the interactions in Table 6.6.

Table 6.5. Main effects

Variable	Value	B	Wald	df	Probability	Exp(B)
Frequency	n/a	0.066	1.376	1	0.241	1.068
Group	n/a	n/a	12.482	2	0.02	n/a
	grammar controls	0.580	3.119	1	0.077	1.786
	vocabulary controls	1.352	12.221	1	<0.001	3.866
Verb type	irregular	0.730	3.627	1	0.057	2.076
Verb-end complexity	n/a	n/a	8.346	2	0.015	n/a
	VV-D	0.918	8.283	1	0.004	2.504
	VC-D	0.340	1.377	1	0.241	1.405

There is no significant main effect of frequency. The main effect of group is significant. Both control groups perform better than the G-SLI group, and this difference is approaching significance for the grammar controls and is highly significant for the vocabulary controls. The main effect of verb type is marginally significant, with irregulars being easier than regulars. There is a significant main effect of verb-end complexity, and a significant difference between performance on VVC-D verbs (the reference category) and VV-D verbs, with the latter being easier. There is no significant difference between VVC-D and VC-D verbs, although the latter are numerically easier.

Table 6.6. Interactions

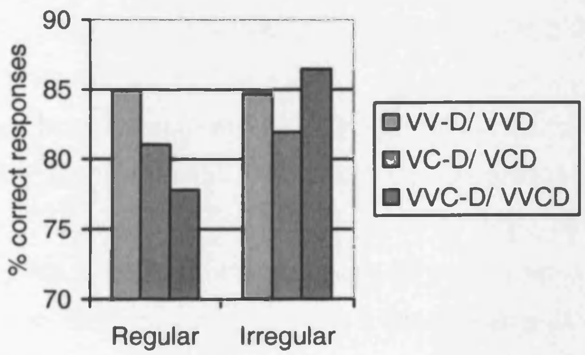
Variables	Value	B	Wald	df	Significance	Exp(B)
Group x verb type x verb-end complexity	n/a	n/a	4.199	4	0.380	n/a
Group x verb type	n/a	n/a	2.609	2	0.271	n/a
Group x verb-end complexity	n/a	n/a	3.549	4	0.470	n/a
Verb type x verb-end complexity	n/a	n/a	0.6220	2	0.045	n/a

None of the interactions involving group are significant. However, there is a significant interaction between verb type and verb-end complexity, which indicates that verb-end complexity affects regulars and irregulars differently. The source of this interaction is shown in Table 6.7 and Figure 6.3. It can be seen that while performance on regular verbs decreases with increasing complexity, this is not the case for irregulars.

Table 6.7. % correct responses for each condition, collapsed across participant groups

Condition		Mean (SD)
Regular	**VV-D**	84.96 (35.32)
	VC-D	81.02 (38.85)
	VVC-D	77.78 (40.32)
Irregular	**VV-D**	84.72 (35.71)
	VC-D	81.94 (37.95)
	VVC-D	86.46 (34.24)

Figure 6.3. % correct responses for each condition, collapsed across participant groups



This interaction is unpacked by investigating whether for each participant group verb-end complexity impacts on the judgement of regulars and irregulars. Recall that verb-end complexity is only predicted to impact on regular verbs for the G-SLI group. The main effects and interactions are presented for the G-SLI group in Table 6.8.

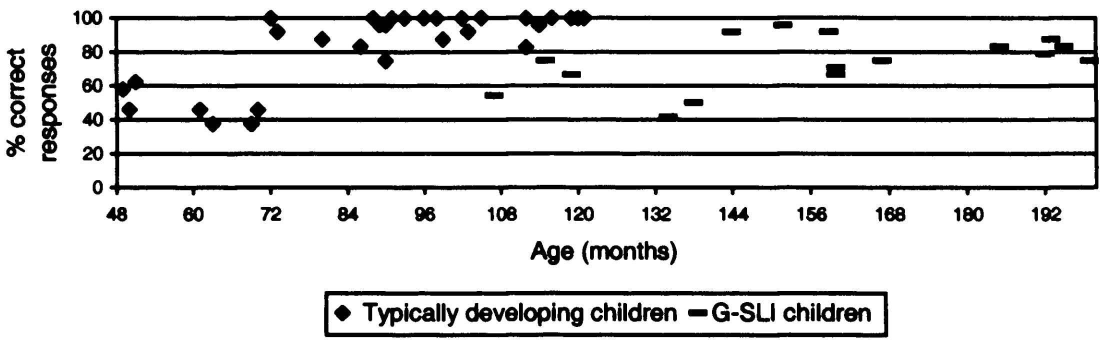
Table 6.8. Main effects and interactions for the G-SLI group

Variable	Value	B	Wald	df	Significance	Exp(B)
Frequency	n/a	0.087	0.968	1	0.345	1.091
Verb-end complexity	n/a	n/a	7.907	2	0.019	n/a
	VV-D	0.952	7.905	1	0.005	2.590
	VC-D	0.330	1.286	1	0.257	1.391
Verb type	irregular	0.670	2.430	1	0.119	1.955
Verb type x verb-end complexity	n/a	n/a	6.288	2	0.043	n/a

For the G-SLI group, verb type and frequency have no significant effect on performance, but verb-end complexity, as predicted, does. The Exp(B) values indicate that as verb complexity increases, performance decreases. This difference is significant for VVC-D versus VV-D verbs, but not for VVC-D versus VC-D verbs. Furthermore, there is a significant interaction between verb-end complexity and verb type, meaning that verb-end complexity impacts differently on regular and irregular performance, again as predicted. The same analysis carried out for each control group did not reveal any significant main effects.

The next question of interest is how individual G-SLI children's performance on regular verbs compares to that of typically developing children. The score for individual children's performance on regulars is shown plotted against age in Figure 6.4.

Figure 6.4. % correct responses for individuals for regular verbs, plotted against age



Chance performance is at 34-66% (binomial, $p < 0.05$), and I consider scores of 90% and higher to be at ceiling levels. Typically developing children score at chance until between four and six years of age. However, from six years onwards the majority score at ceiling. It is noticeable that this transition in performance from chance to ceiling is abrupt, and that very few children score between 66% and 90%. These results are consistent with those of Rice, Wexler, Marquis & Hershberger (2000), who found that irregular verbs can be modelled linearly, but that the model for regular verbs includes a phase of accelerated growth. Note that variation occurs in the typically developing population, with several children not achieving ceiling scores at even eight or nine years of age.

G-SLI children show a range of scores. Three score at chance levels (RP, BD and CT). Three score above 90% (LJ, DD, MS), although it is notable that not a single child achieves a perfect score, in stark contrast to the younger typically developing children. For the ten children who achieve scores between 66% and 90%, I argue that phonological complexity is one of the factors affecting performance.

A further issue concerns the optionality of children's responses. Recall that each verb is used three times in this task. How many children get a particular verb correct on 0/3 occasions (i.e. are consistently wrong), 3/3 occasions (i.e. are consistently right) or on 1/3 or 2/3 occasions (i.e. are optionally right)? The data in Table 6.9 show that it is very rare for a G-SLI child to always choose the unmarked form of the regular verb, and in each case this happens with a child whose performance on the task as a whole is at chance, indicating that he can reasonably be interpreted as just guessing the answer. Interestingly, it is more common for a G-SLI child to make 0/3 correct responses for an irregular verb, and this can occur with children who score above chance on the task overall. For example, two children invariably choose *hold* as the past tense form of *to hold*, even though their overall performance on irregulars is 87.5% and 83.33% respectively. A third child always chooses *ride* as the past tense form of *to ride* even though his overall performance on irregulars is 87.5%. These errors could reasonably be interpreted either as lexical errors (i.e. having past tense meaning stored in the wrong lexical entry for the wrong form of the verb) or a potential strategy of 'identify the past tense form by the presence of a final /*t*/ or /*d*/' failing in cases where both the past tense form and the stem form end in /*d*/.

Table 6.9. Optionality of G-SLI responses for regular and irregular verbs

Number of correct responses for a particular verb	Regular verbs (%)	Irregular verbs (%)
0	2.34	6.25
1	14.06	9.38
2	42.19	24.22
3	41.41	28.91

6.4. Discussion

6.4.1. Summary of results

The predictions for this study were that prosodic complexity at the inflected verb-end would:-

- Impact on G-SLI children's judgement of regular but not irregular verbs
- Not affect the control groups' performance on either regular or irregular verbs.

These predictions are indeed borne out by the data. For G-SLI children, as the verb-end complexity of regular verbs increases, judgement becomes less accurate. G-SLI children's judgement of irregular verbs is not affected by complexity. In contrast to the G-SLI group, verb-end complexity does not affect the typically developing groups' performance on regular verbs. These results were predicted from the finding that G-SLI children have impoverished representations of non-words containing complex syllabic structures (Chapter 5). Judging whether a regular verb is in the past tense requires a representation of verb-end complexity. Complexity does not significantly influence the judgement of irregular verbs, because that task does not require a representation of verb-end complexity. If a child has a representation of, for example, **sol* as past tense of *sold*, then it is still possible to differentiate stem (*sell*) and past forms because of the change in vowel quality, but if he has **roll* as his past tense representation of *rolled* then he won't be able to distinguish the past from the stem form. Although, for the G-SLI group, the presence of a verb cluster reduces judgement accuracy, nuclear complexity before the cluster has no significant effect on performance.

G-SLI children perform worse than both control groups on the task as a whole. This difference is not quite significant for the grammar control group, but highly significant for the vocabulary control group. Although verb-end complexity did not affect performance in either control group, it would be worth investigating at a future date the impact of prosodic complexity on judgements in younger typically developing children, e.g. up to age of 6;06. It would also be worth investigating whether phonological factors are responsible for lowering the performance of those older children who don't yet perform at ceiling.

6.4.2. The impact of prosodic complexity on past tense judgements

The results reported in this chapter are consistent with van der Lely and Ullman's (2001) hypothesis that children with G-SLI store both regular and irregular past tense verbs. If such children have impaired prosodic representations, it will be more difficult for them to store/create complex forms, and inflection is more likely to be omitted in production/judgement tasks. I claim that G-SLI children's pattern of performance on the judgement task reported in this chapter is due to a representational deficit, but could they not instead have a perceptual deficit? Any auditory processing account would contend that a /d/ is less salient after a consonant (e.g. in *sold*) than after a vowel (e.g. in *played*).

In practice it is difficult to tease apart perceptual and representational deficits, a point which has received scant attention in the psycholinguistic literature. Relevant to this issue is work by Dupoux and colleagues (Dupoux, Kakeli, Hinose, Pallier & Mehler, 1999). They found that adult Japanese listeners perceive 'illusory' vowels inside consonant clusters in VCCV stimuli: they hear a stimulus such as 'ebzo' as 'ebuzo'. They also have difficulty in discriminating between VCCV and VCuCV stimuli such as 'ebzo' and 'ebuzo'. The authors interpret this as being due to the reduced syllable inventory of Japanese, which disallows word-medial clusters. French listeners, whose language allows word-medial clusters, do not hear an illusory vowel inside clusters, and have no difficulty in discriminating VCCV from VCuCV. However, they do have problems discriminating pairs of stimuli that differ in vowel length, such as 'ebuzo' and 'ebuuzo', presumably because vowel length is not contrastive in French. What these results show is that the adult imposes the phonological structure of his native language upon novel perceptual stimuli. The problem is not a perceptual one as such - there is no claim that these adults have a perceptual deficit. Rather the difficulty is one of impoverished phonological representations influencing perception.

It is plausible that something similar is happening in the case of G-SLI children. Their difficulty in judging inflection on verbs with complex verb-ends could result not from a perceptual difficulty, but from a difficulty in representing complex phonological structures. Given evidence from other sources that children with G-SLI do not suffer from a deficit in auditory perception (van der Lely, Rosen & Adlard, in press), I argue that prosodic complexity impacts on judgements of regular past tense inflection because of difficulties in representing complex verb ends.

Does a difficulty in representing prosodic complexity at the verb end impact on the performance of all the G-SLI children who participated in this study? Child BD's performance is relevant here: BD performs at chance on the judgement task, even though his chronological age-appropriate performance on the TOPhS (see Chapter 5.3.2.2) indicates normal phonological representations. Additional linguistic data indicate that BD's chance performance is likely to be part of a more widespread deficit in morphosyntax. BD's spontaneous speech (unpublished data, Hilary Gardner, personal communication, December 2002) reveals that he lacks basic knowledge of the morphosyntactic properties of inflectional suffixes, i.e. what stems they attach to. Examples of his errors (highlighted in red) are shown below:-

- **Angelas know_ the way* – instead of *Angela knows the way.*
- **And hes like_ crisps* – instead of *He likes crisps.*
- **Started to hurting* – instead of *Started to hurt.*
- **He_ asking the children to singing the songs* – instead of *He's asking (or he asks) the children to sing the songs.*

His performance on the Word Structure subtest of the Clinical Evaluation of Language Fundamentals (CELF-III^{UK}, Semel, Wiig & Secord, 2000), administered at the same time as the judgement task, reveals a very low score (z-score < -2.33), indicating problems with morphology as a whole. Examples of his errors are shown below:-

- **Here the baby _ felling asleep* – instead of *Here the baby is sleeping* or *Here the baby is falling asleep.*
- **He is called a sings* – instead of *He is called a singer.*
- **I want some of toys* – instead of *I want some of those.*
- **Soon he will slidings* – instead of *Soon he will slide.*
- **Today he is walking and tomorrow he will walks* – instead of *Tomorrow he will walk.*

The **Angelas know the way* error seen in BD's spontaneous language production has also been reported in typically developing 2 year olds (Thornton & Tesan, 2003). It is proposed that a particular parameter, the Infl-Type parameter, characterises languages as having either featural or affixal inflection (Lasnik, 1995). In English, the parameter setting is affixal for lexical verbs, meaning that inflection is merged with the verb in morphology (*knows*) and not with the syntactic component (**Angelas*). However, Thornton and Tesan find that in elicited production experiments, some children appear to hypothesise that English has featural inflection, e.g. **Hes fit in there* and **The bears like cheese* (instead of *The bear likes cheese*). Thornton and Tesan interpret their results as indicating that 2 year old children are trying out different hypotheses about the morphology of English. This could also be an explanation for BD's errors.

Elena Gavruseva (2003) has a different interpretation of similar data. She analyses the *-s* not as the 3rd person agreement marker but as the contracted form of the auxiliary *is*. The overuse of *-s* coincides with a period where the child makes omissions of *-ing* in obligatory contexts. According to Gavruseva, the child makes these errors because he has not yet mastered the aspectual features of the auxiliary. It is clear from both BD's spontaneous and elicited data that he does not yet use *-ing* appropriately. Whatever the analysis of this particular construction, his performance both on elicited tasks of morphology and in natural conversation reveals that he has severe problems of morphosyntax, and that poor performance on the judgement task is due to those problems and not to difficulties with phonology. No perceptual deficit account of SLI predicts such gross morphosyntactic errors.

In conclusion, the G-SLI group as a whole has difficulty in representing prosodic complexity, and this impacts on their judgement of past tense forms – judgements of regulars with verb-end clusters are less accurate than those of verbs with no cluster. Two points are important though – not every G-SLI child has problems with prosodic complexity, and a child can have difficulty with past tense judgements despite normal phonology. This suggests that a model whereby phonological deficits are the cause of morphological deficits (e.g. Joanisse, 2004; Joanisse & Seidenberg, 1998, 2003) cannot hold for all children. Instead, I argue for a model whereby G-SLI children have deficits in syntax and morphology (by definition, as these are the criteria on which they are selected), and many have additional phonological deficits which impact on morphology. The precise manner in which phonological complexity impacts on past tense inflection will be characterised further in Chapters 7 and 8, and its impact on other types of inflection and on derivational morphology will investigated in Chapters 9, 10 and 11. I use the findings from these studies to develop a model of the impairment in G-SLI which van der Lely, myself and colleagues term the Deficit in Computational Grammatical Complexity (CGC) hypothesis.

Chapter 7. The impact of verb-end complexity on past tense formation

7.1. Introduction

7.1.1. Chapter outline

Chapter 6 revealed that prosodic complexity impacts on G-SLI children's grammaticality judgements of inflection, with verbs whose inflection creates clusters being harder to judge as inflected than those without clusters. The work in this chapter investigates whether this effect of prosodic complexity is evident in production too.

In Section 7.1.2 I justify my reasons for investigating the effects of prosodic complexity in a production task. In Section 7.2 I present the method and in Section 7.3 the results. I summarise the results in Section 7.4.1, and in Section 7.4.2 I discuss how prosodic complexity affects inflection in both typically developing children and those with G-SLI.

7.1.2. Verb-end complexity

The aim of the study reported in this chapter is to compare the ability of typically developing children and children with G-SLI to inflect verbs whose inflected verb end contains no cluster, a two-consonant cluster and a three-consonant cluster.

In Chapter 6 I showed that there is a significant difference in performance on a judgement task between inflected verb ends that have a cluster and those that don't. Two questions remain: (1) Will this effect of verb-end complexity also be evident in production? and (2) Will three-consonant clusters be harder than two-consonant clusters in production? Certainly three-consonant clusters are acquired later than two-consonant clusters both word-initially (Kirk, 2003; Smit, 1993) and word-finally (Kirk, 2003). Kirk demonstrates that a particular three-consonant cluster is acquired only after both constituent two-consonant clusters are acquired. For example, if a child can produce a word-final nasal + stop cluster (e.g. ***/pɪŋk/***) and a word-final stop + */s/* cluster (e.g. ***/baks/***), then he can produce a word-final nasal + stop + */s/* cluster (e.g. ***/lɪŋks/***). If, however, one of those two-consonant clusters is reduced, then the three-consonant cluster will be reduced in the same way. For example, if ***/pɪnk/*** is reduced to ***/pɪk/***, then ***/lɪŋks/*** will be reduced to ***/lɪks/***. Therefore, in the tasks reported in this chapter, I contrast two- and three-consonant clusters. VC-D and VVC-D verbs are not contrasted, because the judgement task reported in Chapter 6 found that although performance was numerically higher on VC-D verbs, this difference was not statistically significant.

7.2. Method

7.2.1. Verb stimuli

Four conditions, with 8 verbs in each condition, were selected. Note that the VC-D verbs are those used in Chapter 4. The data for all conditions were collected as part of the same study and are analysed altogether here. For the purposes of the analysis presented here, the results for the VC-D legal and the VC-D illegal verbs are combined. The characteristics of each verb group with regards to verb-end complexity and phonotactics are presented in Table 7.1, and the full set of verbs is listed in Appendix C.1. It proved impossible to balance the four conditions for past tense frequency given the constraints on which stimuli could be used – not only did the verbs have to be familiar to children of the ages taking part in the experiment, but they had to enter into stimulus sentences that were syntactically correct and pragmatically plausible. However, if a correlation is found between frequency and performance, then frequency can be partialled out of the analysis.

Table 7.1. Phonological characteristics of stimuli

Condition	**Consonants at verb end**	**Cluster legal?**	**Examples**	**Frequency**	
				CO-B*	**F&K****
VV-D	1	n/a	*sewed, poured*	1.502	2.419
VC-D legal	2	✓	*killed, wrapped*	1.647	2.089
VC-D illegal	2	x	*touched, robbed*	1.175	1.425
VCC-D	3	x	*solved, punched*	0.721	1.133

* Frequencies obtained from COBUILD, CELEX database; ** Frequencies obtained from Francis & Kucera

7.2.2. Procedure: Elicitation task

The procedure was has already been described in Section 4.2.2. There are 4 practice items using irregular verbs, and 32 experimental items which are listed in Appendix C.2. One pseudo-randomised list was created for all participants.

7.2.3. Participants

Particpants are the same as those who took part in the study in Chapter 4, and whose details can be found in Section 4.2.3. For ease of reference, details of these participants are presented in Table 7.2. Recall that the LA1 control group provides a grammar age match for the G-SLI group. In terms of vocabulary ability, the G-SLI group falls between the LA1 and LA2 groups.

Table 7.2. Participant details

Measure		**G-SLI** N = 14	**LA1 controls** N = 14	**LA2 controls** N = 14
Age	**Mean**	12;03	6;00	9;06
	Range	9;09 – 16;08	4;06 – 7;05	7;06 – 12;00
TROG	**Raw, mean**	12.86	10.76	16.43
	Raw, range	6 – 17	6 – 16	12 – 19
	z-score, mean	-1.67	-0.14	0.12
BPVS	**Raw, mean**	79.93	60.00	94.21
	Raw, range	47 – 104	33 – 81	69 – 120
	z-score, mean	-1.67	0.28	0.28

7.2.4. Predictions

Predictions for the elicitation task are different for the G-SLI and the control groups. For the G-SLI group I predict that accuracy will decrease as cluster complexity increases. For typically developing children I predict no effect of cluster complexity, although if there is I expect it to be in the same direction as for the G-SLI group, i.e. as complexity increases, accuracy will decrease. The majority of errors are predicted to be bare stem errors.

7.2.5. Coding of responses

Responses were coded as follows:-

- Correct — correct inflection, e.g. *hop* → *hopped*
- Bare stem — inflection missing, e.g. *rob* → *rob*
- Stem-final consonant deletion — (inflection supplied, stem-final consonant omitted), e.g. *milk* → *milt*
- Other responses — e.g. *wash* → *washing*; *judge* → *jumped*

On the rare occasions when a child corrected himself, the first response was accepted for analysis.

7.3. Results

7.3.1. Correct responses

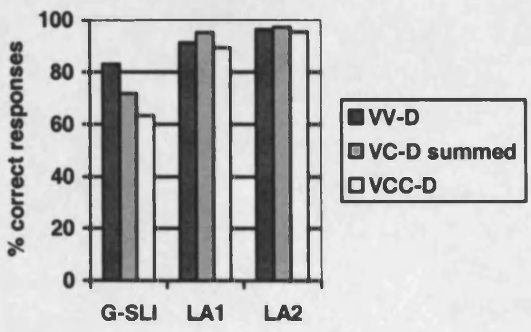
Data from one participant in the LA1 group were discarded for two reasons. Firstly, her answers were at times muffled and therefore difficult to transcribe accurately. Secondly, assuming that her answers have been scored correctly, she achieved a score of 28.13%, which is 3.01 standard deviations below the mean for her group. Her data are therefore unreliable and not representative of the LA1 group as a whole.

Correct responses to the elicitation task are shown in Table 7.3 and Figure 7.1. I include the results for the sum of the VC-D legal and illegal stimuli, i.e. all two-consonant clusters, and the total percentage of correct responses for each group.

Table 7.3. % correct responses

Condition		G-SLI (N=14)	LA1 (N=13)	LA2 (N=14)
VV-D	Mean (SD)	83.04 (20.57)	91.35 (10.69)	96.43 (7.64)
VC-D summed	Mean (SD)	71.88 (30.40)	95.19 (8.90)	97.32 (6.81)
VCC-D	Mean (SD)	63.39 (37.82)	89.42 (11.23)	95.54 (7.92)
Total	Mean (SD)	72.54 (28.54)	92.79 (6.92)	95.54 (5.81)

Figure 7.1. % correct responses



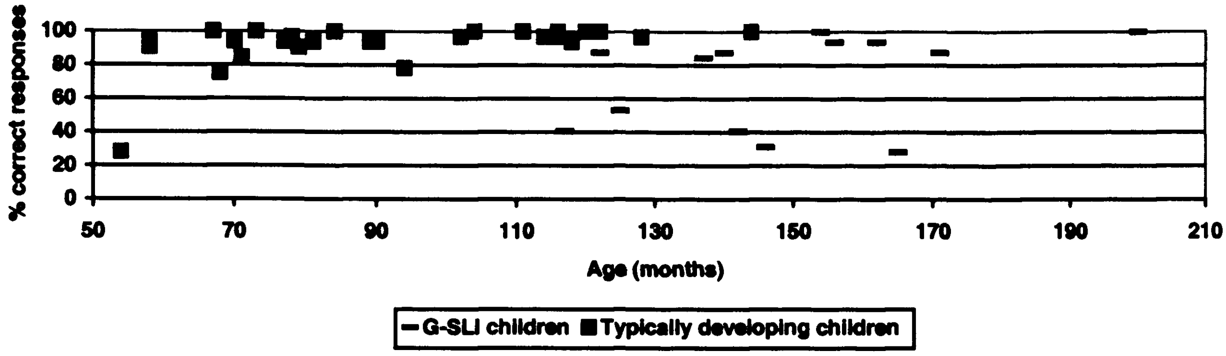
Correct responses were analysed using a 3 (Group: G-SLI, LA1, LA2) x 3 (Condition: VV-D, VC-D, VCC-D) ANOVA. This revealed significant main effects of group, $F\ (2, 38) = 7.450$, $p = 0.002$, and condition, $F\ (2, 38) = 5.498$, $p = 0.006$. The interaction between group and condition was also significant, $F\ (2, 38) = 3.860$, $p = 0.007$.

The group x condition interaction was investigated with a series of pair-wise comparisons between conditions within each participant group. For the G-SLI group, performance is significantly better for the VV-D compared to the VC-D condition, $t(13) = 2.253$, $p = 0.042$, for the VV-D compared to the VCC-D condition, $t(13) = 2.797$, $p = 0.015$, and for the VC-D compared to the VCC-D condition, $t(13) = 2.267$, $p = 0.041$. For the LA1 and LA2 groups, however, none of the pairwise comparisons reach significance. The pattern of correct performance with respect to verb-end complexity is therefore VV-D > VC-D > VCC-D for the G-SLI group, but VV-D = VC-D = VCC-D for the controls.

The interaction was further investigated using a series of one-way ANOVAs within each condition, in order to identify group differences. For the VV-D condition the effect of group is marginally significant, $F(2, 40) = 3.184$, $p = 0.053$. Post hoc comparisons (Bonferroni-corrected) showed that, importantly, the G-SLI group do not differ from the LA1 controls on this set of verbs, $p = 0.408$, but are marginally worse than the LA2 controls, $p = 0.051$. The two control groups do not differ from one another, $p = 1.000$. For the VC-D condition the effect of group is significant, $F(2, 40) = 7.745$, $p = 0.002$, with the G-SLI group performing worse than the LA1 controls, $p = 0.008$ and the LA2 controls, $p = 0.003$, but the two control groups performing equivalently, $p = 1.000$. For the VCC-D condition the effect of group is likewise significant, $F(2, 40) = 7.371$, $p = 0.002$. Again, the G-SLI group perform worse than the LA1 controls, $p = 0.019$, and the LA2 controls, $p = 0.003$, but the two control groups do not perform differently from one another, $p = 1.000$. The G-SLI group therefore perform worse than the control groups on all conditions, except for the VV-D conditions, where they perform equivalently to the LA1 group.

As was the case for the TOPhS results reported in Section 5.3.1, there is a wide range of scores within the G-SLI group, and a wide variety of ages within the G-SLI and control groups. The performance of each child (including the outlier from the LA1 control group, whose data are excluded from the statistical analysis) is plotted against age in Figure 7.2. For the G-SLI group there is no significant correlation between age and score on the elicitation task. For the LA1 and LA2 controls, however, the correlation is significant, $r = 0.543$, $p = 0.046$ and $r = 0.538$, $p = 0.048$ respectively.

Figure 7.2. Individual performance plotted against age.



The major error type for all three participant groups is the bare stem form of the verb. Bare stem forms account for the following proportion of total errors: G-SLI, 73.17%, LA1, 80.00%, LA2, 93.38%. Table 7.4 shows the proportion of bare stem errors expressed as a percentage of the total number of responses for each cell.

Table 7.4. % Bare stem errors

Condition	G-SLI	LA1	LA2
VV-D	8.93 (15.83)	6.73 (12.97)	2.68 (7.24)
VC-D summed	21.88 (25.79)	4.33 (9.40)	2.68 (6.93)
VCC-D	27.68 (30.29)	7.69 (9.60)	4.46 (7.92)

So few errors were made in the 'stem-final consonant deletion' and 'other responses' categories that I combine these in Table 7.5. Deletion of a stem-final consonant is rare, and is confined to VCC-D clusters. The examples are: *solved* → */sɔld/, *milked* → */mɪlt/, *banged* → */bænd/. The G-SLI group makes this error only 4 times, the LA1 controls only twice, and the LA2 controls not even once.

One G-SLI child, OD, makes 11 responses with *-ing* (omitting the auxiliary), e.g. **Yesterday I laughing* instead of *Yesterday I laughed.* This type of error is not made by any of the control children. The remainder of the errors made by the G-SLI and LA1 groups are semantic/phonological in nature, e.g. **waved* for *weighed*, **saved* for *solved* and **sign* for *sighed.* The only error amongst the LA2 controls is an irregularisation, whereby */suːd/ is produced in lieu of *sewed.*

Table 7.5. 'Stem-final consonant deletion' and 'other response' errors combined

Condition		G-SLI (N=14)	LA1 controls (N=13)	LA2 controls (N=14)
VV-D	**Mean (SD)**	8.04 (14.38)	1.92 (6.93)	0.89 (3.34)
VC-D summed	**Mean (SD)**	6.25 (12.57)	0.48 (1.74)	0.00 (0.00)
VCC-D	**Mean (SD)**	8.93 (15.05)	2.98 (5.48)	0.00 (0.00)
All	**Mean (SD)**	7.37 (13.64)	1.47 (3.97)	0.22 (0.84)

7.4. Discussion

7.4.1. Summary of results

The elicitation task was designed to investigate the impact of verb-end prosodic complexity on inflection, a factor predicted to affect rates of past tense inflection in G-SLI children but not in language-matched controls. This was indeed found to be the case: for the G-SLI group, accuracy decreases as the prosodic complexity of the inflected verb end increases, with performance being most accurate on verbs without a cluster and least accurate on verbs with a three-consonant cluster. The overwhelming error type for all stimuli is the bare stem form. Another error occurs for verbs ending in a three-consonant cluster, whereby deletion of the stem-final consonant is attested in G-SLI and LA1 children, but only at low rates. These results are in line with the findings of the grammaticality judgement task in Chapter 6, whereby an increase in prosodic complexity led to decreased judgment accuracy.

Note the lack of correlation between performance and age, and the lack of mastery for the elicitation task amongst the G-SLI children: only 5 out of the group of 14 achieve a score greater than 90%. Rice (2004) reports similar results: she and her colleagues found that in their group of children with SLI, even those as old as 14 asymptote at lower levels than typically developing children, suggesting that they never fully master tense marking. The point that I will argue in Section 7.4.2 is that for the G-SLI group, some of this lack of mastery can be accounted for by difficulties with phonology.

7.4.2. The impact of phonological complexity on past tense inflection

The work in this chapter confirms that G-SLI children have difficulty with regular past tense inflection when it introduces clusters at the word-end. Although coronal stop deletion is a phonologically driven phenomenon in English, stimuli were constructed so as to minimise its effects, by ensuring that the child produced the verb as part of a sentence whereby the word following the verb began with a vowel. This is further evidence that suffix omissions should be interpreted as morphological rather than phonological errors. However, the rate of suffix omission is affected by phonological complexity.

Most errors occur on verbs whose inflected verb end contains a three-consonant cluster. Out of the two types of errors that are possible for VCC-D stimuli, the overwhelming majority are bare stem forms: there are only a few examples of the stem-final consonant being deleted, e.g. *milked* → */mɪlt/*. The consonant in this stem-final position is probably deleted for perceptual reasons (e.g. see discussion in Cote, 2002). An important cue to stops is their release burst, which is weak before another consonant. In other words, the weak release of, for example, the */k/* in *milked*, would explain why it is deleted. It is perhaps surprising then that stem-final consonant deletion in these three-consonant clusters is not more frequent. It must surely point to the privileged status of the stem, whereby the pressures to retain the stem are stronger than the pressures to retain the suffix. That this is the case for the G-SLI children just as it is for typically developing chidren is further evidence against a perceptual deficit causing SLI, and in favour of the morphological deficit argued for in Chapters 3 and 4.

As an aside, it is probably these same perceptual difficulties that mean that clusters such as */lkt/* and */ŋd/* are not found word-finally (with the exception of the rarely used word *mulct*). Interestingly, Stemberger and Bernhardt (1997) claim that an error such as *milked* → */mɪlt/* cannot occur, although Bernhardt and Stemberger (1998: 484) produce data from a child who reduces the cluster in *jumped* to */mt/*. This child also produced the form **jumpted*, presumably as a way of syllabifying the */t/* as an onset rather than as a word-final consonant. Errors such as **jumpted* are not attested in any of the children who participated in the task reported here.

Recall that in Chapters 3 and 4 I showed that G-SLI children performed worse on illegal compared to legal VC-D verbs. A hypothesis that unites the twin impact of phonotactics and phonological complexity on inflection is that G-SLI children are impaired in forming morphologically complex forms. They therefore have to rely on the storage of such forms and/or their creation by analogy. I propose that G-SLI children store inflected verbs with complex verb ends less accurately, and are less likely to create these forms by analogy, because of their deficit in representing phonological complexity. Within the set of verbs whose inflected verb-end is a cluster, those with illegal clusters are produced less accurately because illegal clusters are less frequent than legal clusters. Compared to irregular verbs, regulars are more likely to have verb-end clusters, and only regular verbs have illegal clusters. These differences in the phonology of regular and irregular verbs could contribute to G-SLI children's greater than expected impairment on regular verbs as compared to irregulars: performance on regular forms in both elicitation and judgement tasks is less accurate than predicted from these children's general language age.

So far, models of past tense inflection have failed to take the phonological complexity of the verb end into account, but the results presented in this chapter strongly suggest that they should. Verb-end complexity can be incorporated into a single mechanism account, where past tense is handled by mappings between phonological and semantic content of past tense forms. In this case, the phonological deficit would cause the morphological deficit. In a dual mechanism model whereby language is a modular system, and where morphology and phonology are separate modules, it is possible that difficulties with phonology could impact on the output from the morphological module. The phonological deficit would be distinct from the morphological deficit, and not causally linked to it. Pinker and Ullman (2002) do not address the impact of phonology on inflection in their otherwise thorough defence of the Words and Rules model. The only phonological factor they mention is 'phonotactic naturalness' (p.472), but they do not specify what they mean by this.

In conclusion, the evidence points to G-SLI children having independent deficits in syntax, morphology and phonology which all impact on tense. Some questions are still outstanding in relation to the interaction between phonology and past tense morphology: the use of the /*ɪd*/ allomorph, and phonological effects on the production of irregular forms. These issues will be the focus of the next chapter.

Chapter 8. The syllabic allomorph /ɪd/

8.1. Introduction

8.1.1. Chapter outline

The work reported in Chapter 7 revealed that the prosodic complexity of the inflected verb end affects regular past tense production in children with G-SLI. The work reported in this chapter investigates two further ways in which phonology is predicted to impact on past tense formation. I investigate inflection rates when the syllabic allomorph /ɪd/ is required, and I consider whether the phonological characteristics of irregular verb stems affect the production of bare stem forms and over-regularisations.

The outline of the chapter is as follows. In Section 8.1.2 I discuss studies that have considered the acquisition of /ɪd/ in typically developing children and those with SLI. In Section 8.2 I present the method and in Section 8.3 the results. In Section 8.4.1 I summarise the results, and in Section 8.4.2 I discuss the different ways in which phonology affects past tense inflection, based on the studies in Chapters 6, 7 and 8. Finally, in Section 8.4.3, I develop a model whereby independent deficits in syntax, morphology and phonology impact on tense in children with G-SLI.

8.1.2. The /ɪd/ allomorph

The /ɪd/ allomorph is added to verb stems that end in an alveolar stop, i.e. /t/ or /d/. Because of the limited phonological environment in which it can occur, it is rarer than the other two past tense allomorphs, /t/ and /d/, with a frequency of 22.6% of all past tense tokens (Stemberger & MacWhinney, 1986b). It has been recognised since Berko's pioneering elicitation studies of typically developing children (Berko, 1958) that of the three past tense allomorphs, /ɪd/ is the last to be acquired. There is, of course, a certain amount of variability. For example, Bernhardt and Stemberger report a child who used /ɪd/ before she was able to create clusters at the inflected verb end, acquiring the past tense in the sequence VV-D (e.g. *cried*) → /ɪd/ (e.g. *needed*) → VC-D (e.g. *kissed*) (Bernhardt & Stemberger, 1998). In adult speech, /ɪd/ is omitted more frequently than other allomorphs, both in natural speech (Stemberger, 1983) and under experimental conditions (Stemberger & MacWhinney, 1986b).

Is there any indication that /ɪd/ is problematic for children with SLI? Oetting and Horohov (1997) found that 6-year old SLI children produced lower rates of suffixation for regular verbs taking /ɪd/ compared to those taking /t/ and /d/ inflection, but only when compared to chronological age-matched controls; there were no differences compared to language-matched controls. However, the authors had only 4 /ɪd/ verbs in their study.

Marchman, Wulfeck and Weismer (1999) looked at past tense marking in regular and irregular verbs in children with SLI (aged 6;01-12;00) and chronological age-matched controls. Not surprisingly, the SLI group scored significantly lower than the controls. The two groups showed a different pattern of behaviour with regards to regulars and irregulars ending in alveolar consonants (i.e. /t/ and /d/) versus those ending in non-alveolar consonants.

- For regular verbs, the SLI group made more bare stem errors for stems ending in /t/ and /d/ than for stems ending in another consonant, whereas for the controls there was no such difference.
- For irregulars, the SLI group made fewer over-regularisations and more bare stem errors for stems ending in /t/ or /d/ than for stems ending in another consonant. Again, for the controls there was no such difference.

Taken together, these results show that SLI children have difficulty using the /ɪd/ suffix.

Van der Lely and Ullman's (2001) study did not include any regular verbs ending in /t/ or /d/. However, it is logical to predict that if /ɪd/ is harder for typically developing children, young SLI children and adults, then it will also be problematic for individuals with G-SLI.

Why should /ɪd/ be so difficult? An influential hypothesis, the 'affix-checking' hypothesis (Berko, 1958; MacWhinney, 1978), claims that children analyse verbs that end in /t/ or /d/ as already being marked for the past tense: children do not add a suffix because the word already appears to end in one. A related hypothesis (Bybee & Slobin, 1982; Taatgen & Anderson, 2002) is that many irregular stems (35% by type) end in /t/ or /d/, and either have identical past tense forms (e.g. *hit, put, bet*) or have just a change of vowel in their past tense form (e.g. *hide/hid, ride/rode, meet/met*). By contrast, only 11% of regular stems end in /t/ or /d/. It is conceivable that the child might analyse regular stems that end in /t/ or /d/ as past irregular forms. Rumelhart and McClelland (1986), Pinker and Prince (1988) and Marchman, Wulfeck and Weismer (1999) suggest that producing bare stem forms for stems ending in /t/ and /d/ might be an over-generalisation of the no-change pattern of verbs such as *hit* and *hurt*.

An alternative possibility is that children use /ɪd/ less frequently because doing so changes the metrical structure of the verb. Perhaps their phonology resists adding a syllable to a word because they prefer to maintain the word's metrical structure. Or perhaps two-syllable forms more generally are inflected at lower rates. We can try to tease apart whether it is the change in metrical structure or the creation of a two-syllable inflected form that is the difficulty in the following way: we can compare two-syllable stems with monosyllabic regulars that end in /t/ and /d/, to test whether a two-syllable output is generally harder than a one-syllable output. Even if we discount the possibility that a two-syllable inflected output is more difficult to produce than a one syllable inflected output, this still doesn't tell us whether the issue is the change in metrical stucture or analysis of the stem as being already inflected. This is because the allomorphs /t/ and /d/ are the same sounds as the stem-final consonants that require /ɪd/. It is difficult to know how to tease these two factors apart at this stage, although in Chapter 9 I present data from plural inflection (where the stem-final consonants that require /ɪz/ are not only /s/ and /z/, but also /tʃ/, /ʃ/, /dʒ/ and /ʒ/) which suggest that a change in metrical structure is at least partly a factor.

The study presented here also considers the over-regularisation of irregular verbs. There are abundant references that over-regularisations are common in child phonology (Berko, 1958; Ervin, 1964; Kuckjaz, 1977; Marcus, Pinker, Ullman, Hollander, Rosen & Xu, 1992) and that they also occur in SLI (Marchman *et al.*, 1999; van der Lely & Ullman, 2001). How are over-regularisations of irregular forms related to suffixations of regular forms? We would expect the same phonological pressures that are relevant for regular suffixation to be relevant for the suffixation of irregulars. In Chapter 7 I showed that regular verbs which when inflected have the form VV-D are easier to inflect than VC-D forms. We would predict this to also be true for irregulars, with more overregularisations for VV stems, e.g. *fly*, than for VC stems, e.g. *dig*. What about stems that would take the /ɪd/ allomorph if they were over-regularised, such as *fight*, *rode*? Given the previous studies discussed in this section, we would predict such stems to be over-regularised less frequently than those ending in VV. In other words, the constraints on the output of over-regularisation of irregular verbs should be the same as those on output of inflection of regular verbs: we would expect over-regularisation to be less common when it produces a cluster or requires /ɪd/ compared to when it produces no cluster and does not require /ɪd/.

8.2. Method

8.2.1. Verb stimuli

6 conditions, with 8 verbs in each condition, were selected for the elicitation task. 3 conditions consisted of regular verbs, and 3 of irregular verbs. The phonological and morphological characteristics of these conditions are presented in Table 8.1. For the full list of stimuli, see Appendix D.1. Although in previous experiments I have endeavoured to choose stimuli with simplex rather than complex onsets, it has not been possible to do so for the irregular verbs used here because not enough irregular verbs exist that not only fit into the 3 conditions, but that are also matched on frequency. As I explained in Section 6.2.1, it is very difficult to match regular and irregular verbs for frequency. Although frequency was not a significant factor in that particular study, I was more careful to match all conditions for frequency in this study. I did so by calculating the frequency for regular verbs over the entire verbal paradigm, e.g. for *rent* this is the frequency of *rent*, *rents*, *rented* and *renting*, and for irregular verbs using the frequency of the past tense form only. On these measures, regular and irregular verbs are matched for frequency.

Table 8.1. Verb stimuli and their characteristics

Condition	Morphology	Phonological characteristics	Examples	Mean frequency*
d/t-ɪd	Regular	Ends in an alveolar stop	*rent, start*	3.920
VV-D	Regular	Ends in a long vowel	*sew, tie*	3.815
SS-D	Regular	2 syllables, no final consonant	*follow, whisper*	3.935
d/t-ɪd	Irregular	Ends in an alveolar stop	*ride, bite*	3.769
VV-D	Irregular	Ends in a long vowel	*fly, draw*	3.860
VC-D	Irregular	Ends in a non-alveolar consonant	*choose, break*	3.846

* Frequency obtained from Francis and Kucera

For the regular verbs, there are no stimuli which when inflected have clusters at the verb end, because such verbs have already been investigated in Chapter 7. For irregular verbs, children with G-SLI add the suffix to both present and past tense forms, though more commonly to the present form (van der Lely, unpublished data). The irregular stimuli used here are chosen so that both their present and past tense forms have the same phonological characteristics: this makes the calculation of over-regularisation rates more straightforward.

When choosing irregular stimuli, there are several phonological characteristics that deserve to be looked at in more detail. For example, competition between vowels of the two forms of vowel-change irregulars, such as *fall* and *fell*, is known to affect over-regularisation rates in spontaneous child language (Stemberger, 1993). Dominant vowels are those with features such as [+high], [+low], [+back] and [+round]. Verbs with dominant vowels in the stem form and non-dominant vowels in the past tense form (e.g. *fall/fell*) are more likely to be over-regularised. The stimuli that I use here, however, are chosen to extend my work on verb-end phonology. For verbs ending in */t/* or */d/*, only those with a vowel change, rather than no-change verbs (e.g. *hit*, *cut*), were selected, so that correct tense-marking can be distinguished from bare stem production. The terminology that I use for labelling the irregular stimuli might strike the reader as being a little unusual because they are labelled by the incorrect form of the target. It might seem strange to think of *throw* as being labelled VV-D, but my aim is to make parallels with the regular stimuli, e.g. *sewed*, and to show that the same phonological pressures are predicted to impact on the inflection of both sets of verbs.

8.2.2. Procedure

An elicitation task was used which was slightly modified from that used in Chapters 4 and 7. In that task the child was primed with both the stem and past tense form in an effort to improve on the low scores that the G-SLI children achieved in van der Lely and Ullman's (2001) task. From the results of the task used in Chapters 4 and 7, I felt confident that the children would achieve scores higher than in van der Lely and Ullman's task without the need for the past tense to be presented. I also dispensed with the toy animals used to introduce the task in Chapters 4 and 7, judging that the youngest typically developing children would be able to manage the task without them, while the older G-SLI children would consider the toys, and therefore the task, too 'babyish'.

The lead in was of the form '*Everyday I get a present. Yesterday I _____ a present*'. There were 2 practice sentences, using irregular *go* and *have*, neither of whose past tense corresponds to the phonological characteristics of the irregular experimental stimuli. One randomised, set order was created for all participants, and is presented in Appendix D.2.

8.2.3. Participants

In this experiment, and in those reported in Chapter 9, I use a different method for selecting control groups. In the experiment reported in Chapter 6 I used individual matches. While this has the advantage of ensuring that each child in the G-SLI group is matched on exact or near-exact raw scores of grammar and vocabulary measures, the disadvantage is that one cannot see any pattern of development within the control children unless one afterwards assigns them to groups based on age (the method I use in the studies reported in Chapters 4, 5, 7, 10 and 11). For this experiment and those described in Chapter 9, I use three control groups, each encompassing approximately a twelve-month age band. 36 control children are divided into 3 groups of 12 each.

- LA1 controls – aged 5;04-6;06 (mean age 6;00)
- LA2 controls – aged 6;07-7;06 (mean age 7;01)
- LA3 controls – aged 7;09-8;05 (mean age 8;02)

As was the case for the control groups in previous chapters, two receptive language tests were administered – the TROG and the BPVS. Scores for all participant groups are presented in Table 8.2.

Table 8.2. Participant details

Measure		**G-SLI** N = 13	**LA1 controls** N = 12	**LA2 controls** N = 12	**LA3 controls** N = 12
Age	**Mean**	13;05	6;00	7;01	8;02
	Range	9;08 – 17;09	5;04 – 6;06	6;07 – 7;06	7;09 – 8;05
TROG	**Raw, mean**	12.62	14.53	16.17	17.17
	Raw, range	6 – 17	12 – 17	14 – 19	15 – 19
	z-score, mean	-1.76	0.59	0.59	0.47
BPVS	**Raw, mean**	77.23	68.80	76.92	91.00
	Raw, range	47 – 107	60 – 81	63 – 97	71 – 106
	z-score, mean	-1.80	0.55	0.35	0.49

In order to determine how the G-SLI group compares with each of the control groups on each of the language measures, a series of independent samples t-tests was carried out. For the TROG, the G-SLI group did not differ from the LA1 controls, t (23) = -1.876, p = 0.073, but it scored significantly worse than the LA2 group, t (23) = -3.335, p = 0.003, and the LA3 group, t (23) = -4.750, p < 0.001. Even though the G-SLI and LA1 groups are not well-matched, the LA1 group is the closest control group to the G-SLI group in terms of grammatical ability. For the BPVS, the G-SLI group did not score significantly differently to the LA1 group, t (23) = 1.722, p = 0.105, or the LA2 group, t (23) = 0.057, p = 0.995, but did score worse than the LA3 group, t (23) = -13.7692, p = 0.033. Therefore the

LA2 group is the best-matched control group for vocabulary. These matches are indicated in red in Table 8.2. The LA3 group has significantly higher grammatical and vocabulary abilities than the G-SLI group, and is included in order to determine the pattern of typical development.

8.2.4. Predictions

Predictions for both regular and irregular verbs rely on the hypothesis that stems ending in *-d/t* are less likely to be suffixed than those ending in VV. For regular verbs, I predict that the G-SLI and control groups will both achieve lower rates of inflection with ***d/t-ɪd*** verbs compared to VV-D verbs. No studies have previously compared one- and two-syllable verbs. The SS-D verbs chosen for this study have trochaic foot structure and end in a vowel, so they are not any more prosodically complex than monosyllabic VV-D verbs. I therefore predict that performance will be equivalent on VV-D and SS-D verbs for both groups, with ***d/t-ɪd*** verbs harder than both.

For irregular verbs, interest centres on two types of errors: over-regularisations (e.g. *throw* → **throwed*) and bare stem responses (e.g. *meet* → **meet*). For the G-SLI group, I predict lower rates of over-regularisation for ***d/t-ɪd*** and VC-D verbs compared to VV-D verbs. Because the factors that reduce the likelihood of over-regularisation are the same that promote bare stem responses, I predict more bare stem responses for ***d/t-ɪd*** and VC-D verbs than for VV-D verbs. I make no predictions as to whether this group's performance will be lower for ***d/t-ɪd*** or VC-D verbs. For typically developing children, I have shown that prosodic complexity does not affect regular inflection (see Chapter 7), and so I predict no distinction between VV-D and VC-D verbs for either over-regularisation or bare stem responses, but I predict fewer regularisations and more bare stem responses for ***d/t-ɪd*** verbs compared to both VV-D and VC-D verbs.

8.2.5 Coding of responses

Responses to regular verbs are coded as follows:-

- Correct, e.g. *whisper* → *whispered*
- Bare stem, e.g. *lift* → *lift*
- Other, e.g. *sew* → */suːd/*, *chew* → *cheweded*, *weigh* → *waved*, *weigh* → *did weigh*, *lift* → *am lifted*, no response

Responses for irregulars were coded somewhat differently. Over-regularisation errors cannot by definition apply to a regular verb stem, since past tense suffixation is the correct response. Two types of over-regularisation are possible for irregulars – one on a present stem, e.g. **throwed*, and one on a past stem, e.g. **wonned*. Both types are counted together here. Responses for irregular verbs were coded as follows:-

- Correct, e.g. *steal → stole*
- Bare stem, e.g. *meet → meet*
- Over-regularisation, e.g. *throw → throwed, win → wonned*
- Other, e.g. *get → had, tear → did tear, steal → stolen, blow → blows, dig → dag,* no response

8.3. Results

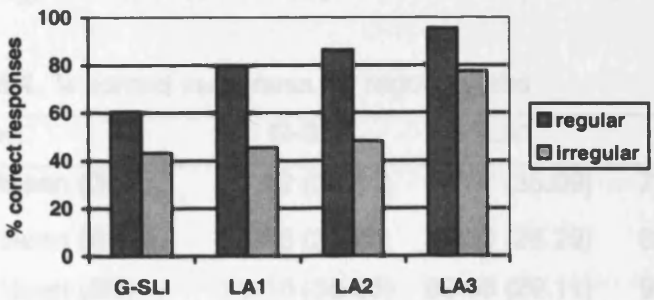
8.3.1. Comparing performance on regular and irregular verbs

In this section I first compare regular and irregular performance. I next investigate the impact of phonology within regular, and then within irregular, verbs. The percentage of correct responses for regulars and irregulars within each group is set out in Table 8.3.

Table 8.3. % correct responses for regular and irregular verbs

Condition		G-SLI	LA1	LA2	LA3
Regular	Mean (SD)	60.58 (32.03)	79.86 (27.57)	86.46 (18.13)	95.49 (4.15)
Irregular	Mean (SD)	43.27 (40.09)	45.49 (20.68)	48.26 (23.20)	77.43 (19.34)

Figure 8.1. % correct responses for regular and irregular verbs



A 4 (Group: G-SLI, LA1, LA2, LA3) x 2 (Condition: regular, irregular) ANOVA revealed significant main effects of group, $F(3, 45) = 4.960$, $p = 0.005$, and of condition, $F(1, 45) = 68.000$, $p < 0.001$, and a near-significant interaction between the two, $F(3, 45) = 2.758$, $p = 0.053$. Post hoc multiple comparisons (Bonferroni-corrected) revealed that the

G-SLI group only score significantly lower than the LA3 controls, p = 0.003. No other pair-wise comparisons reached even marginal significance.

The interaction between group and condition was investigated by carrying out paired samples t-tests comparing regular and irregular performance within each group. For the G-SLI children the difference in performance does not reach signifcance, t (12) = 1.967, p = 0.073. However, for each of the control groups, performance is significantly higher for the regular condition: LA1, t (11) = 6.766, p < 0.001, LA2, t (11) = 6.587, p < 0.001, LA3, t (11) = 3.430, p = 0.006. The interaction between group and condition was further investigated with a series of one-way ANOVAS to reveal group differences. A one way ANOVA within the regular condition reveals a significant effect of group, F (3, 45) = 5.074, p = 0.004. Post hoc multiple comparisons (Bonferroni-corrected) reveal that the G-SLI group performs significantly worse than the LA2 group, p = 0.048, and the LA3 group, p = 0.003, but not significantly worse than the LA1 group. None of the pairwise comparisons between the control groups were significant. For the irregular condition, the analysis reveals a group effect, F (3, 45) = 4.114, p = 0.012. Pair-wise comparisons revealed that this time the G-SLI group performs significantly worse than only the LA3 controls, p = 0.020, who had superior language abilities. The LA1 group also performs significantly worse than the LA3 controls, p = 0.040, but no other pairwise comparisons are significant. The interaction between group and condition therefore arises from particularly poor performance by the G-SLI group on regular verbs.

8.3.2. Performance on regular verbs

To investigate performance within just the regular verbs, I first consider the impact of phonology on response accuracy. The data are presented in Table 8.4.

Table 8.4. % correct responses for regular verbs

Condition		G-SLI	LA1	LA2	LA3
t/d-ɪd	**Mean (SD)**	51.92 (38.81)	66.67 (35.09)	73.96 (30.83)	91.67 (11.10)
VV-D	**Mean (SD)**	58.65 (31.20)	83.33 (26.29)	88.54 (18.81)	97.92 (4.87)
SS-D	**Mean (SD)**	71.15 (35.86)	89.58 (29.11)	96.88 (10.83)	96.88 (7.77)

A 4 (Group: G-SLI, LA1, LA2, LA3) x 3 (Condition: *t/d-ɪd*, VV-D, SS-D) ANOVA revealed significant main effects of group, F (3, 45) = 5.074, p = 0.004, and of condition, F (2, 44) = 17.255, p < 0.001, but no significant interaction between group and condition. Post hoc comparisons (Bonferroni-corrected) revealed that performance across groups is significantly worse for the ***t/d-ɪd*** compared to the VV-D condition, t (48) = -0.3.061, p = 0.004, worse for the ***t/d-ɪd*** compared to SS-D condition, t (48) = -5.694, p < 0.001, and worse for the VV-D compared to the SS-D condition, t (48) = -2.873, p = 0.006. Therefore the overall pattern of performance is ***t/d-ɪd*** < VV-D < SS-D, with all groups responding to phonology in a similar way.

Of the different error types for regular verbs, the one of principle interest is the bare stem response, as this is the error that characterises G-SLI children's past tense productions. Bare stem errors account for the majority of errors in all groups and for all stimulus types. They account for the following proportion of total errors: G-SLI, 75.59%, LA1, 98.24%, LA2, 100%, LA3, 69.23%. Table 8.5 shows the proportion of bare stem errors expressed as a percentage of the total number of responses for each cell.

Table 8.5. % Bare stem errors for regular verbs

Condition		**G-SLI**	**LA1**	**LA2**	**LA3**
t/d-ɪd	**Mean (SD)**	43.27 (38.06)	31.25 (35.25)	26.04 (30.83)	7.29 (11.25)
VV-D	**Mean (SD)**	24.04 (33.67)	16.67 (26.29)	11.46 (18.81)	1.04 (3.61)
SS-D	**Mean (SD)**	25.00 (36.08)	10.42 (29.11)	3.13 (10.83)	1.04 (3.61)

A 4 (Group: G-SLI, LA1, LA2, LA3) x 3 (Condition: ***t/d-ɪd***, VV-D, SS-D) ANOVA on the bare stem response scores revealed significant main effects of condition, F (2, 44) = 23.880, p < 0.001, but no significant effect of group, F (3, 45) = 2.630, p = 0.062, and no interaction between group and condition, F (6, 45) = 1.610, p = 0.153. Paired samples t-tests (Bonferroni-corrected) revealed significantly more bare stem errors for the ***t/d-ɪd*** compared to the VV-D condition, t (48) = 4.706, p < 0.001, and for the ***t/d-ɪd*** compared to the SS-D condition, t (48) = -5.655, p < 0.001, but no significant difference between the VV-D and SS-D conditions, t (48) = 1.268, p = 0.211. Therefore the overall pattern of bare stem errors is ***t/d-ɪd*** > VV-D = SS-D, with all groups responding to phonology in a similar way.

This pattern of results for correct and bare stem responses raises the issue of why correct performance for the VV-D condition is lower than that for the SS-D condition, and yet the number of bare stem errors is no different. Presumably another type(s) of error is being made on the VV-D condition which lowers correct performance on these verbs. This is indeed the case. Some children with G-SLI (4/13 children) double mark verbs in the VV-D condition, producing forms such as **paided* and **chewed**ed*. 2 produce irregularisations for VV-D verbs. For example one child (GS) irregularises *sew*, *tie* and *row* as ***suː*, ***tuː* and ***ruː* respectively.

It should be noted that children from all three control groups also make occasional irregularisations, and not just on the VV-D condition, for example *shoot* → **shote*, *ride* → **rid* and *dig* → **dag*. A child from the G-SLI group, (QC) makes confusions with other morphological suffixes, producing forms such as **am lifted*, **am rent* and **answering* (without the auxiliary). In one of his answers he puts the past tense suffix on the direct object, producing **am marrying a dancered*. The errors illustrated here are not made by *any* of the typically developing children participating in the study.

8.3.3. Performance on irregular verbs

In this section I consider performance on irregular verbs, first investigating correct responses, and then errors. Table 8.6 shows the proportion of correct responses for each group according to condition.

Table 8.6. % correct responses on irregulars

Condition		G-SLI	LA1	LA2	LA3
t/d-ɪd	**Mean (SD)**	41.35 (41.26)	42.71 (27.42)	50.00 (26.11)	80.21 (18.81)
VV-D	**Mean (SD)**	39.42 (39.48)	38.54 (17.24)	39.58 (24.91)	64.58 (26.02)
VC-D	**Mean (SD)**	49.04 (42.84)	55.21 (27.93)	55.21 (28.43)	87.50 (15.99)

A 4 (Group: G-SLI, LA1, LA2, LA3) x 3 (Condition: ***t/d-ɪd***, VV-D, VC-D) ANOVA revealed significant main effects of group, $F(3, 45) = 4.114$, $p = 0.012$, and of condition, $F(2, 44) = 17.993$, $p < 0.001$, but no significant interaction between group and condition. Post hoc t-tests (Bonferroni-corrected) revealed that performance across groups is significantly better for the ***t/d-ɪd*** compared to the VV-D condition, $t(48) = 2.562$, $p = 0.014$, worse for the ***t/d-ɪd*** compared to the VC-D condition, $t(48) = -3.810$, $p < 0.001$, and worse for the VV-D compared to the VC-D condition, $t(48) = -5.600$, $p < 0.001$. Therefore the overall pattern of performance is VV-D < ***t/d-ɪd*** < VC-D, with all groups responding to phonology in a similar way.

Next I investigate the types of errors made. Two types of response - bare stems (e.g. *meet* → **meet*) and over-regularisations (e.g. *throw* → **throwed*) - are of particular interest because of the predicted impact of phonology. Recall that the prediction was for more bare stem responses to occur for ***t/d-ɪd*** verbs, and for more over-regularisations to occur for VV-D verbs. I consider bare stem responses first. The proportion of such responses for each group and each condition is set out in Table 8.7.

Table 8.7. % bare stem responses for irregulars

Condition		**G-SLI**	**LA1**	**LA2**	**LA3**
t/d-ɪd	**Mean (SD)**	33.65 (40.66)	22.92 (34.47)	16.67 (28.37)	2.08 (7.22)
VV-D	**Mean (SD)**	24.04 (33.79)	9.38 (22.06)	11.46 (22.27)	2.08 (4.87)
VC-D	**Mean (SD)**	25.96 (31.23)	11.46 (23.36)	12.50 (21.32)	1.04 (3.61)

To investigate the impact of group and condition on bare stem response rates, a 4 (Group: G-SLI, LA1, LA2, LA3) x 3 (Condition: ***t/d-ɪd***, VV-D, VC-D) ANOVA was carried out. This revealed a significant main effect of condition, $F(3, 45) = 6.003$, $p = 0.004$, but the effect of group was not significant, $F(3, 45) = 2.346$, $p = 0.085$. The two way interaction was not significant. The main effect of condition was further investigated by a series of paired samples t-tests (Bonferroni-corrected) on the pooled data for all participant groups. There are significant differences between the ***t/d-ɪd*** and VV-D conditions, $t(48) = 2.739$, $p = 0.009$, and between the ***t/d-ɪd*** and VC-D conditions, $t(48) = 2.558$, $p = 0.014$, but not between the VV-D and VC-D conditions, $t(48) = -0.843$, $p = 0.404$. This indicates that there are more bare stem responses for verbs in the ***t/d-ɪd*** condition than for those in either the VV-D or VC-D conditions. In other words, the number of bare stem errors is ***t/d-ɪd*** > VV-D = VC-D, with all groups responding in the same way.

Next I investigate the impact of condition on over-regularisation errors. Recall that the prediction was for most over-regularisations on the VV-D condition. The proportion of over-regularisations for each group and each condition are presented in Table 8.8.

Table 8.8. % over-regularisation errors for irregulars

Condition		**G-SLI**	**LA1**	**LA2**	**LA3**
t/d-ɪd	**Mean (SD)**	15.38 (24.56)	32.29 (27.42)	31.25 (27.95)	13.54 (14.56)
VV-D	**Mean (SD)**	28.85 (33.22)	51.04 (17.24)	48.96 (26.36)	33.33 (26.83)
VC-D	**Mean (SD)**	17.31 (28.66)	31.25 (18.84)	29.17 (25.75)	11.46 (14.56)

A 4 (Group: G-SLI, LA1, LA2, LA3) x 3 (Condition: ***t/d-ɪd***, VV-D, VC-D) ANOVA revealed significant main effects of group, $F(3, 45) = 2.907$, $p = 0.045$, and of condition, $F(3, 45) = 22.691$, $p < 0.001$, but no significant interaction. Post hoc multiple comparisons

(Bonferroni-corrected) revealed that no one group produced more over-regularisations than any other. Paired samples t-tests comparing the effects of condition found that verbs in the VV-D condition are over-regularised significantly more often than those in the ***t/d-ɪd***, t (48) = 4.658, p < 0.001, and VC-D conditions, t (48) = 6.698, p < 0.001, but that there was no significant difference between the ***t/d-ɪd*** and VC-D conditions. For over-regularisations, then, the pattern is VV-D > ***t/d-ɪd*** = VC-D

The next question of interest is to investigate the relative pattern of bare stem and over-regularisation responses for each group. The figures in Tables 8.7 and 8.8 strongly suggest that the error pattern for typically developing children is for more over-regularisations than bare stem responses. T-tests comparing the number of bare stem and over-regularisation errors for each group reveal that there is no significant difference in the two types of errors for the G-SLI group, t (12) = -0.542, p = 0.598. However, for the LA1 and LA2 groups, there is a marginally significant difference in error production, with over-regularisations more common than bare stem responses, t (11) = 2.127, p = 0.057, and t (11) = 2.146, p = 0.055 respectively. For the LA3 group the difference is highly significant, t (11) = 3.559, p = 0.004. These data indicate that although the total number of correct responses for irregular verbs does not differ between the G-SLI group and the LA1 and LA2 control groups (see Section 8.3.1), the pattern of error responses is different.

Bare stem responses and over-regularisation errors are the major error types for all participant groups. All other categories of errors put together make up 8.34% of the G-SLI group, and only 1.73%, 1.74% and 1.40% of responses for the LA1, LA2 and LA3 groups respectively. All groups make occasional over-regularisations on past tense stems, e.g. **stoled*, lexical substitutions, e.g. **had* for *got*, and incorrect irregularisations, e.g. **dag* for *dug*. There is only one example of a past participle form being produced by a control child (**chosen*), whereas 3 G-SLI children make this error a total of 8 times. One type of error that is made by some G-SLI children, but never by any of the control children, is best described as revealing impaired morphosyntactic knowledge. Examples of this error involve variously adding the past tense suffix either to the object of the verb, e.g. **steal a watched* (Child HD); using a different form of the correct verb, e.g. **am lead* (Child QC), and adding the wrong suffix to the object of the verb, e.g. **speeching* (instead of *gave a speech*, Child SL).

8.4. Discussion

8.4.1. Summary of results

On the elicitation task reported in this chapter, children with G-SLI show no statistically reliable difference in performance between regular and irregular verbs, whereas all three control groups show an advantage for regulars. For regular verbs as a whole, G-SLI children perform equivalently only to their grammar-matched controls, whereas for irregulars they perform equivalently to their grammar and vocabulary controls. These results confirm van der Lely and Ullman's (2001) findings that the G-SLI group lack the regularity advantage shown by typically developing children, and in comparison to controls, perform relatively worse on regulars.

For regular verbs, the G-SLI and the typically developing groups respond to the phonological characteristics of the inflected verb end in the same way. All groups are less accurate in inflecting verbs in the *t/d-ɪd* condition compared to those in the VV-D condition. Correct performance is highest for the SS-D condition. The most frequent error for each group and condition is the bare stem response.

For irregular verbs, there are more bare stem responses for verbs in the *t/d-ɪd* condition than for those in either the VV-D or VC-D conditions. The VV-D condition is over-regularised more frequently than the *t/d-ɪd* or VC-D condition, but there was no difference in levels of over-regularisation for *t/d-ɪd* and VC-D. Although the correct level of performance of the G-SLI group and the two younger control groups on irregular verbs is indistinguishable, the errors they make are different. The control children make more over-regularisation errors than bare stem responses, whereas for the G-SLI group there is no significant difference in error type, although bare stem responses are more numerous.

The finding of lower over-regularisation rates for the irregular VC-D condition mirrors the findings in Chapter 7 of lower inflection rates for the regular VC-D condition, while the results for regular and irregular verbs indicate that the use of the /*ɪd*/ allomorph is disfavoured for some reason. Hence we can conclude that the effects of phonological complexity cut across morphology type.

8.4.2. The range of phonological effects on past tense inflection

In Chapters 6, 7 and 8 of this thesis I have demonstrated that a range of phonological factors affects regular and irregular past tense inflection in both typically developing children and those with G-SLI. In this section I discuss the results of the study reported in the present chapter, which concern suffixation of the /*ɪd*/ allomorph in regulars and the impact of verb-end phonology on bare stem and over-regularisation errors in irregular verbs. I also broaden the discussion to include the results of the studies in Chapters 6 and 7, which were concerned with the impact of verb-end prosodic complexity on regular inflection, in preparation for the model of a Deficit in Computational Grammatical Complexity that I propose in Section 2.4.3. As a starting point for the discussion, I summarise the results of the relevant studies. In Table 8.10 I present a summary of the results with respect to regular verbs. Here, comparisons are made on the basis of correct performance. In red are the results that differ between groups, and in blue are the results that are identical for both groups.

Table 8.10. Summary of results for regular verbs

Phonological factor	Task	Condition	G-SLI group's performance	Control groups' performance
Prosodic complexity	Judgement	VV-D vs. VC-D vs. VVC-D	VV-D > VC-D > VVC-D	VV-D = VC-D = VVC-D
Prosodic complexity	Elicitation	VV-D vs. VC-D vs. VCC-D	VV-D > VC-D > VCC-D	VV-D = VC-D = VCC-D
/ɪd/ allomorph	Elicitation	VV-D vs. SS-D vs. *t/d-ɪd*	SS-D > VV-D > *t/d-ɪd*	SS-D > VV-D > *t/d-ɪd*

The results for the studies of irregular verbs are summarised in Table 8.11 and are presented slightly differently to the results for regulars. For the judgement task (Chapter 6), comparisons are made on the basis of correct performance. For the elicitation task (Chapter 8), comparisons are made on the basis of bare stem and over-regularisation errors. Again, results that differ between groups are highlighted in red, and results that are identical for both groups are highlighted in blue.

Table 8.11. Summary of results for irregular verbs

Phonological factor	Task	Condition	G-SLI group's performance	Control groups' performance
Prosodic complexity	Judgement	VV-D vs. VC-D vs. VVC-D	VV-D = VC-D = VVC-D	VV-D = VC-D = VVC-D
/ɪd/ allomorph	Elicitation (bare stem)	VV-D vs. VC-D vs. *t/d-ɪd*	VV-D = VC-D > *t/d-ɪd*	VV-D = VC-D > *t/d-ɪd*
	Elicitation (over-reg.)	VV-D vs. VC-D vs. *t/d-ɪd*	VV-D > VC-D = *t/d-ɪd*	VV-D > VC-D = *t/d-ɪd*

The data clearly show that the use of the /ɪd/ allomorph is problematic not only for those children with G-SLI, but also for their typically developing, language-matched controls aged 5;04-8;05. In Section 8.1.2 I hypothesised that one reason why /ɪd/ might be difficult to use is that children resist changing the metrical structure of words. It is impossible to distinguish between this hypothesis and the affix-checking hypothesis because the past tense allomorphs /t/ and /d/ are the very stem-final sounds that condition the selection of the /ɪd/ allomorph. However, the data collected here allow us to rule out one possibility – that a two-syllable inflected output is generally more difficult to produce than a one-syllable output, and that suffixation with /ɪd/ is disfavoured for this reason. A comparison of inflection rates between two-syllable verbs and one-syllable verbs ending in /t/ or /d/ reveals fewer bare stem errors for the SS-D compared to the *t/d-ɪd* condition. That indicates that the lower inflection rate for the *t/d-ɪd* condition is either due to the change in metrical structure being disfavoured or due to the segmental content of the stem making the verb appear already inflected. An interesting result is that there are more correct responses for the SS-D than for the VV-D condition. One possible explanation for this finding is that all two-syllable verbs with strong-weak stress (i.e. the stimuli used in this study) are regular, and therefore a sw pattern is an unambiguous cue to morphology. Note that there are a handful of two-syllable irregular verbs, but the majority are prefixed and have final primary stress, e.g. *begin*, *become*, *withhold*, *upset* etc.

As regards irregular verbs, I agree with Stemberger and Middleton (2003) that, with cognitive science's fixation on the differences between regular and irregular verbs, differences within irregulars have been largely ignored. (A notable exception is Bybee & Slobin, 1982). In this study I have shown that the phonological characteristics of irregulars do indeed affect morphological behaviour. As predicted, there are more bare stem responses for the *t/d-ɪd* condition than for either the VV-D or VC-D conditions, and verbs in the VV-D condition are over-regularised more frequently than those in the *t/d-ɪd* or VC-D conditions.

Both G-SLI and control groups double mark irregulars, e.g. **flewed*, **tored*, **wonned*, etc., albeit on only a few occasions. Presumably this reflects the fact that irregular past tense forms are stored in the lexicon, and therefore available for suffixation. Children with G-SLI also double mark regular forms, e.g. **tieded*, **cheweded*, **roweded*, but this type of error is not found in the typically developing children tested here. I tentatively interpret this result as providing a clue that regular past tense forms are stored by children with G-SLI, meaning that they are available as stems for suffixation. Alternatively, the fact that responses of this type are not recorded in the elicitiation of VV-D verbs in the task in Chapter 7 suggests that their presence here is due to priming effects from the ***t/d-ɪd*** verbs that are also used here, e.g. *needed* and *started.* If this is the explanation, then it suggests that representations of regular past tense forms in the G-SLI group are not as secure as those of typically developing children, since the typically developing children resist priming.

When trying to understand the phonological pressures on the over-regularisation of irregular verbs we have to distinguish two separate pressures – what an ideal output should look like, and how difficult the change from input to output is. *Ride* already has the shape of an ideal output, whereas *tear* does not. This could explain why verbs such as *tear* are over-regularised more frequently than verbs such as *ride* (e.g. Bybee & Slobin, 1982). A second reason why *tear* might be over-regularised more often than *ride* is that adding a /*d*/ to a stem ending in a vowel ending is easier in some way than adding an extra syllable. In terms of syllabic structure, there is nothing more complex about **rided* than there is about **teared. *Rided* has two syllables, but it is still a trochaic foot, but no reason why that should be a problem. The difficulty is in telling which of these two pressures is operative – and indeed both may be.

It would have been informative to investigate more thoroughly the impact of metrical structure on regular inflection. One can imagine that differences in metrical structure could lead to differences in bare stem errors for three-syllable verbs of the form *órganised* versus *remémbered.* However, it would actually prove very difficult to do this, which is why I omitted just such an investigation. The majority of English verbs consist of a single syllable. Two- or three-syllable verbs tend to have lower frequencies and later ages of acquisition, and it would be difficult to find enough stimuli that would be in the vocabulary of our youngest control children.

Now I discuss the effects of verb-end complexity, starting with regular verbs. For the judgement and the elicitation tasks, the results pattern the same way – complexity affects G-SLI children but not the typically developing children tested here. However, I predict that complexity effects would be present at a younger stage of development, because consonant clusters are acquired after singleton consonants: the study would need to look at younger children in order to confirm this. Although for the typically developing children tested here complexity no longer affects performance, use of /*ɪd*/ continues to be problematic.

What about the effect of complexity on irregular verbs? The results are the same for both the G-SLI and the typically developing groups. There is no effect of verb-end complexity on judgement, and I argued in Chapter 6 that this is because children don't need a representation of verb-end complexity when judging whether a form is present or past – the vowel quality indicates which is which. Where it comes to producing bare forms and over-regularised forms though, verb-end complexity does have an effect. When the verb end would contain a cluster if suffixed, we see more bare stem forms and fewer over-regularisations.

Marchman *et al.* (1999) claim that an over-sensitivity to phonology interferes with efficient lexical processing and hence the organisation of general patterns across inflected items. In my view this seems a strange way of looking at the impact of phonology on past tense inflection. I fail to see the logic of why a heightened sensitivity to phonology should cause difficulties with morphology. Isn't it just as likely that, given the cues that phonology provides to morphology, such a sensitivity should aid children with SLI in their acquisition of morphology, causing them to have better morphological skills than their language-matched peers? Nor can I see how this account would tie in with other explanations of SLI as being caused by poor auditory perception, and therefore poor phonology. Furthermore, it is hard to conceive that phonological problems alone can explain all the morphological errors that G-SLI children make. G-SLI children make bizarre morphological errors that are not made by any of the control children tested here. For example, here are some of QC's mistakes from the elicitation task reported in this chapter. In each case the target is the past tense, but he produces a variety of alternative forms:-

- Future e.g. **will get*
- Past with *did* e.g. **did weigh*
- Progressive minus the auxiliary e.g. **tearing*
- Present progressive e.g. **am running*
- Past progressive e.g. **was throwing*
- Auxiliary with bare stem e.g. **am lead*
- Double marking on regulars e.g. **paided*
- Past participle e.g. **broken*
- 3 person singular e.g. **blows*
- Past tense marked on object noun e.g. **am marrying a dancered*

In other words, QC makes ten different error types (in addition to over-regularisations of irregulars), which must surely indicate faulty morphosyntactic knowledge. There are two possibilities, of which he may be using one or other, or both: (1) QC is creating suffixed forms by rule, but is unsure which particular suffix has past tense meaning and what the -*ed* suffix should attach to, or (2) QC is selecting already suffixed forms from his lexicon. Some of these types of errors have been acknowledged by previous studies on SLI but not discussed (e.g. van der Lely & Ullman, 2001), or discussed but left unaccounted for (e.g. Marchman *et al.*, 1999). Nor are they discussed in any connectionist simulations of the past tense deficit in SLI or aphasia. Certainly these errors are problematic for accounts which claim that SLI children have a pattern of normal but delayed development, such as the Extended Optional Infinitive account (Rice, Wexler & Cleave, 1995). However, Bird, Lambon Ralph, Seidenberg, McClelland and Patterson (2003) report instances of adult non-fluent aphasic patients using *-ing* in place of the past tense suffix in elicitation tasks.

In conclusion, it appears that verb-end phonology really does affect past tense morphology in the G-SLI group. The next question of interest is whether phonological factors also affect other types of inflection, and derivational morphology. This is the issue at the heart of Part 3 of this thesis (Chapters 9 to 11). Before moving on, however, I use the next section to develop the CGC hypothesis.

8.4.3. The deficit in Computational Grammatical Complexity (CGC) hypothesis

The CGC hypothesis holds that the linguistic impairment in G-SLI lies in the representation of hierarchical complex structures in three components of grammar – syntax, morphology and phonology. The complex structures that cause difficulty within each component include the following:-

- Syntax – non-local dependencies
- Morphology – concatenation of stem + suffix
- Phonology – consonant clusters and unfooted syllables

Children selected for the G-SLI subgroup have, by definition, difficulties with syntactic structures involving non-local dependencies and problems with inflectional morphology. Many (but not all) also have phonological difficulties.

The CGC hypothesis can account for why tense is affected to such a great extent in the G-SLI subgroup, and across the SLI population more generally. For tense to be realised accurately, the child needs to have mastered complex structures in syntax (V to I movement of tense features), morphology (stem + *-ed*) and phonology (consonant clusters). Deficits in each of those three components will have an additive effect on past tense inflection, resulting in higher levels of suffix omission than are seen in typically developing children of the same general language ability.

A crucial issue in the characterisation of the CGC hypothesis is the nature of structural complexity in each component of grammar. For syntax I assume, following the RDDR hypothesis (van der Lely, 1998), that complexity is defined by non-local structural dependencies. For morphology, I adopt a dual mechanism model (e.g. Pinker, 1999) whereby regulars comprising a stem and suffix are more complex than irregulars, which have no internal structure. For phonology, I adopt a standard model of binary branching structure (e.g. Harris, 1994; see Section 1.2.3.1) whereby branching syllabic structures are more complex than non-branching ones. An important issue concerns whether the hierarchical structure is equivalent across syntax, morphology and phonology. Current linguistic theory suggests that this is not the case – for example, syntax is recursive (noun phrase can occur within another noun phrase), but it is not clear whether prosodic structure is (a syllable cannot occur within another syllable). At present I am agnostic over whether a common impaired algorithm underlies the representation of complexity in each component of grammar, or whether the deficits affecting individual components are independent but highly likely to co-occur. These are issues for further research.

PART 4.

INVESTIGATING THE IMPACT OF A PHONOLOGICAL DEFICIT ON OTHER AREAS OF MORPHOLOGY

Chapter 9. The impact of metrical structure on plural and progressive inflection

9.1. Introduction

9.1.1. Chapter outline

In Chapter 8 I showed that both typically developing children and those with G-SLI omit the past tense suffix more frequently from stems that end in */t/* or */d/* compared to those that end in a vowel. I discussed the possibility that children avoid using the */ɪd/* allomorph for metrical reasons: they resist changing the metrical structure of the verb. This chapter investigates the impact of metrical structure on the use of two further inflectional suffixes – plural *-s* and progressive *-ing*. These types of inflection are claimed by some researchers to be unaffected in SLI (Rice & Wexler, 1996). However, even if G-SLI children have acquired the morphosyntax of plural and progressive inflection, under the proposed model of deficits in different components of grammar (the CGC hypothesis, Section 8.4.3), I predict that phonological complexity will affect the realisation of those inflections.

In Section 9.1.2 I introduce previous studies of plural inflection in typically developing children and children with SLI, and in Section 9.1.3 I do the same for progressive inflection. Section 9.2 is given over to the plural elicitation task. In Section 9.2.1 I present the method, and in Section 9.2.2 the results. Section 9.3 is devoted to the present progressive elicitation task, with the method presented in Section 9.3.1 and the results in Section 9.3.2. In Section 9.4.1 I summarise the results from both studies, and in Section 9.4.2 I discuss the impact of metrical structure on inflection.

9.1.2. The acquisition of plural inflection

Gleitman and Wanner (1982) posit that at first children are not aware that plural forms are composed of two elements, and they do not analyse a plural such as *cats* into its constituent morphemes *cat* and *-s*. The result is that although such children can produce some plural forms, they treat these as single elements of meaning. Alternatively, children at this early stage may use an unmarked stem and instead mark plurality by using a quantifier such as *more* or *two* (Clark, 2003).

Berko (1958) found that the plural is used productively (i.e. to inflect nonsense words) by children as young as four, and that fewer errors are made on plural forms than on past tense forms. Clark claims that English-speaking children acquire the plural so early because English uses just one morpheme to express the concept 'more than one'. Other languages, however, use different morphemes depending on the gender of the noun, whether the noun stem ends in a consonant or a vowel, or whether the noun is preceded by a numeral. It should be harder to express plurality in those languages because there are more forms to learn, and there are conditions on the use of each form (Clark, 2003).

However, Clark oversimplifies the learning task for the English-speaking child. The English plural has three phonologically-conditioned allomorphs - /s/, /z/ and /ɪz/. /ɪz/ is added to sibilant-final stems, i.e. to those that end in /s/, /z/, /tʃ/, /dʒ/, /ʃ/ and /ʒ/. Berko (1958) found that, in the wug task, children were less likely to use /ɪz/ compared to /s/ and /z/, even when adding /s/ and /z/ created word-final clusters. Even children aged 5;06-7;00 inflected nonsense forms such as *gutches* and *nizzes* less than 40% of the time. The reasons for the difficulty with /ɪz/ are hypothesised to be the same as those for /ɪd/, discussed in Section 8.1.2. Certainly, noun stems that end in /s/ or /z/, e.g. *horse* and *rose*, could reasonably be analysed as already being inflected for the plural, although it is less clear that the affix-checking hypothesis holds for stems ending in /tʃ/, /dʒ/, /ʃ/ and /ʒ/. Alternatively, it may be that children don't like to change the metrical structure of the noun. We can discount the possibility that the difficulty is not in producing a two-syllable inflected form more generally by comparing performance on nouns such as *horse* → *horses* with two-syllable nouns which don't change in metrical structure, e.g. *tiger* → *tigers*.

SLI children's acquisition of the plural morpheme is traditionally regarded as less impaired than their acquisition of tense morphemes. Although Leonard, Eyer, Bedore and Grela (1997) found that SLI children aged 3;07-5;09 performed worse than their MLU controls on plural formation, in both spontaneous and elicited data, Oetting and Rice (1993) found no significant differences between SLI children aged 4;07-5;08 and MLU controls on an elicitation task. Why these two studies obtained conflicting findings is not clear, but an analysis of spontaneous plural use collected from Oetting and Rice's SLI group showed high levels (90%) of correct plural use, indicating that plural inflection may be acquired young.

The Surface Hypothesis (e.g. Leonard *et al.*, 1997) claims that inflections that take the form of word-final syllables should not reveal differences between SLI children and their language controls because /ɪz/ is presumably more salient, and is therefore more likely to be identified in the input on a regular basis, giving children more opportunity to hypothesise its grammatical function. Syllabic inflections on monosyllabic nouns can be organised as the weak syllables in strong-weak syllable sequences, thereby increasing the likelihood of successful production. However, the little evidence that is available for plurals does not support the predictions of the surface hypothesis. Oetting and Rice (1993) report worse performance on */ɪz/* for both language-matched controls and SLI children. So */ɪz/*, when compared to the other two plural allomorphs, looks similar to the past tense allomorph */ɪd/* in relation to */t/* and */d/*.

However, even if plural inflection *is* produced at the predicted rate for SLI children's language age, the question remains as to whether output is produced in the same way as by typically developing children. For example, although there were no differences in inflectional accuracy between SLI and control children in Oetting and Rice's study (1993), the authors observed that children in the SLI group often took longer to formulate their responses. However, they provided no data to back up this observation, and because that task was designed to elicit just plural forms, it is not clear whether the SLI children would also have been slower in producing singular forms. The experiment reported in this chapter considers not only inflection rates for the plural affix, but also naming latencies for both singular and plural forms.

Goad (1998) has shown that even when language-impaired members of the KE family (see Section 1.1.2) produce a plural form, that form can have unusual prosodic and melodic characteristics (for example, lack of voicing agreement between the stem and suffix), which suggest that plural formation can occur by means other than normal affixation. Goad hypothesises that the suffix is treated as a stem rather than as an affix, meaning that plural-formation is a process of compounding rather than affixation. She also suggests that some plural forms are stored as unanalysed wholes. SLI subjects therefore have two ways of producing plurals – compounding and memorisation – and this can account for the observation that output can be variously correct and anomalous. However, it should be stressed that the unusual phonological characteristics of plurals produced by language-impaired members of the KE family may be caused by their verbal dyspraxia.

For G-SLI children, plural inflection rates have not been examined in elicitation tasks. However, two studies indicate that *knowledge* of pluralisation is not typical. First, Froud and van der Lely (unpublished manuscript) claim that G-SLI children employ an explicitly learnt strategy for pluralising nouns, on the basis that they give a higher rate of plural responses to mass nouns (e.g. *water* → **waters*) than any of the language-matched control groups. Second, van der Lely and Christian (2000) showed that G-SLI children produce regular plurals inside compounds, e.g **rats-eater*, which are not grammatical in the speech of typically developing children. They claim that this is evidence for G-SLI children storing plural forms. Further evidence for the storage of plural forms comes from frequency effects found in Oetting and Rice's (1993) study: SLI children inflected nouns that are more frequently pluralized in adult speech more successfully than those that are less frequently pluralized, while the control children showed no such difference. It remains to be determined whether G-SLI children also produce */ɪz/* less frequently than other allomorphs, although the prediction is that they will.

9.1.3. The acquisition of progressive inflection

-Ing marks progressive aspect, with aspect being that part of the inflection system that signals whether an event is completed or not, ongoing, iterated etc. *–Ing* is the most regular verb suffix in English (Radford, 1997). It can be attached to the base form of almost any verb, with very few exceptions (e.g. the defective verb *beware*). *–Ing* is one of the first inflections to appear in normal child language (Brown, 1973) and was the verbal inflection used with the highest accuracy in Berko's study (Berko, 1958). Radford analyses *-ing* as being adjoined directly to the verb within the verb phrase and thus not associated directly with any functional category (Radford, 1997). If indeed *-ing* is lexical rather than syntactic, this might explain why it is easily acquired. Alternatively, Berko (1958) suggests that *-ing* is the easiest suffix to acquire because it has only one allomorph, and so is completely phonologically regular. One further reason is that the present progressive is the default construction for expressing the present tense, even when the speaker does not want to communicate that an action is continuous. As such, it is a common construction. *-Ing* forms can also be used with a past tense auxiliary, e.g. *was drinking*.

-Ing is amongst the first verbal inflections to appear in SLI speech, even though it is used without the accompanying auxiliary for a protracted period of time (Cleave & Rice, 1997). Leonard, Deevy, Miller, Charest, Kurtz and Rauf (2003) elicited present and past progressive forms from a group of SLI children, mean age 5;02, a group of language-matched controls and a group of chronological age matched controls. The authors found no differences between the SLI group and the language and age-matched controls on use of *-ing*, and all groups used it more in the present than in the past. Some SLI children who used hardly any auxiliaries still used *-ing*, and this pattern is interpreted as being related to tense – all tense morphemes, including auxiliaries, are known to be susceptible to omission. Montgomery and Leonard (1998) showed that SLI children around the age of 8;06 perform as well as chronological age-matched controls on a grammaticality judgement task involving *-ing*, even though they performed worse relative to their controls on *-ed* and third person singular *-s*. All the stimuli used in the studies discussed here have consisted of just one syllable. To my knowledge, no-one has looked at progressive

formation in verb stems which are longer than one syllable, and which have contrasting stress patterns.

The main aim of the study reported here is to contrast suppliance of *-ing* on verbs with one syllable, those with two syllables and strong-weak stress, and those with two syllables and weak-strong stress. A secondary aim is to consider auxiliary omission. A number of studies indicate that weak syllables are more likely to be retained in strong-weak than in weak-strong configurations, because in the former the weak syllable is part of a trochaic template, the default foot structure in English (e.g. Allen & Hawkins, 1978; Demuth, 1996; Gerken, 1994). Consequently the participants in this study are encouraged to provide a pronoun as the subject of the sentence, in order that the auxiliary be incorporated into the trochaic template, e.g. *{she is} hiding*. Note that for verb stems with weak-strong stress, e.g. *relax*, the auxiliary can still be incorporated into this template, but if it is, then the initial weak syllable of the stem will be outside this template, e.g. *{she is} re{laxing}*. Hence we might expect to find competition between the auxiliary and the initial weak syllable of the stem for inclusion in the template, and more auxiliary omission in constructions involving weak-strong stems than either one syllable or strong-weak stems. However, I stress that investigating auxiliary omission is only a minor aim of this study. Children are likely to respond with a contracted auxiliary, e.g. *he's*, *she's*, and so in these cases, even when the following verb has weak-strong stress, e.g. *relaxing*, metrical effects on auxiliary production are not predicted.

9.2. Plural naming task

9.2.1. Method

9.2.1.1. Noun Stimuli

6 conditions were used, with 12 nouns in each condition. Three conditions were designed to elicit singular nouns and three to elicit plural nouns. These conditions' labels, characteristics and examples of stimuli are presented in Table 9.1. Note that the three-syllable singular stimuli are included for two reasons: (1) to ensure that the number of singular and plural items is balanced, and (2) to provide some slightly more unusual words to keep the children's interest, given that familiar words with high frequencies and low ages of acquisition were necessarily chosen for the other five conditions. These three-syllable items are included in the first analysis (Section 9.2.2.1) when I compare accuracy on singular and plural nouns, but are not included in the naming latency analysis (Section 9.2.2.3) because their form does not match that of any of the plural sets.

None of the plural stimuli have a cluster at the verb end – they end in just a single final consonant, which comprises the suffix. The phonological form of the singular stimuli is also the same – they end in a single consonant. In this way, phonological differences between the stimuli are minimised, making naming latencies easier to compare across conditions. A further reason for keeping the inflected noun end as simple as possible is that plural inflection rates have not been measured in G-SLI children before. I would predict that when inflection introduces clusters, inflection rates will be lower, as they are for the past tense. It makes sense to test in the first instance a basal measure of plural inflection. The main aim of this task is to determine the impact of metrical structure on plurality, and adding conditions with different levels of complexity at the noun-end would have made the number of stimuli too great and the experiment too long. The current design enables us to compare the effects of plurality and metricality without noun-end complexity being a complicating factor. All the stimuli were chosen so that none had a complex onset: given evidence that G-SLI children find complex onsets more difficult than simplex (see Chapter 5), this might affect naming times.

Table 9.1. Noun conditions and their characteristics

Condition*	**Morphological characteristics**	**Phonological characteristics**	**Examples**
s-ø	singular	1 syllable	*dog, cake*
sw-ø	singular	2 syllables, sw stress	*carrot, necklace*
3σ- ø	singular	3 syllables	*tricycle, dinosaur*
s-*z*	plural	1 syllable	*bears, ties*
sw-*z*	plural	2 syllables, sw stress	*anchors, zebras*
s-*ɪz*	plural	1 syllable → 2 syllables	*axes, benches*

* s = stressed syllable, w = weak syllable, ø = no suffix (i.e. the target form is singular), 3σ = 3 syllables, *z* = plural allomorph /*z*/, *ɪz* = plural allomorph /*ɪz*/

9.2.1.2. Procedure

The procedure is a timed naming task, programmed in Visual Basic and presented on a laptop computer. The child is presented with one image at a time to name. Half of these images are designed to elicit the singular form of the noun, and half the images have three objects, e.g. 3 keys, and are designed to elicit the plural form. All the pictures (with one exception, to be discussed later) are taken from Cycowicz, Friedman, Rothstein and Snodgrass (1997). Stimuli are balanced for the complexity of the image and for its familiarity, using the figures in Cycowicz *et al.* (1997). Ages of acquisition were taken from the MRC Psycholinguistic Database (Coltheart, 1981) and anything with a rating greater than 300 (i.e. corresponding to an age of acquisition of 6 and over) was rejected. However, there is a limited choice of s-*ɪz* stimuli, and therefore two words with rating higher than 300 were included - *axes* and *benches*, which are both just over at 311. The figures for familiarity and complexity are shown in Table 9.2. For the full list of stimuli, see Appendix E.1.

Table 9.2. Familiarity and complexity of noun stimuli (Cycowicz *et al.*, 1997)

Condition	**Familiarity**	**Complexity**
s-ø	2.78	2.99
sw-ø	2.62	2.92
s-*z*	2.69	2.96
sw-*z*	2.61	2.71
s-*ɪz*	2.66	2.92

I chose pictures that the children in Cycowicz *et al.*'s (1997) study were the most accurate in naming correctly, i.e. where the 'modal' name matched the target name. There were two exceptions: *lettuce* was named more often as *cabbage* and *church* as *house*, but as the metrical characteristics of both pairs of names are the same, these substitutions do not matter for the purposes of this task. American *turtle* is predicted to be renamed here as *tortoise*. Again, as the two words have the same prosodic structure, it does not matter which the child supplies.

Because of the difficulty in finding suitable s-*ɪz* stimuli, one picture (*purse*) was used which was not part of Cycowicz *et al.*'s set. This picture was taken from the British Picture Vocabulary Scales (Dunn, Dunn, Whetton & Burley, 1997). Obviously it is not ideal to use images from different sets, because they have not been checked for complexity or familiarity, but in this case only one such item has been used. The picture of the purse was altered to be as close in style as possible to Cycowicz *et al.*'s pictures: it does not appear more complex than any of the other pictures, nor is it obviously recognisable as being from a different set.

Half the pictures portrayed a single object, and half had three objects to indicate plurality. 4 different running orders were created, and participants were randomly assigned to each order. There were 6 practice items, 3 singular and 3 plural of metrical structure which were not found in the stimulus set, e.g. *helicopter*, *icecreams*. A total of 72 experimental items, broken into blocks of 24 items, with a short motivating cartoon after each block. The child looked at a fixation point which appeared in the middle of the screen for 500 ms, after which the picture appeared. The picture remained until the child responded, or indicated that he/she did not know the answer. A sound file was recorded for each response, and recording started as soon as the picture appeared on the screen. This sound file was later analysed visually and auditorally in order to determine the naming latency.

9.2.1.3. Participants

The same children participated as in the experiment reported in Chapter 8, with the addition of 2 G-SLI children: SA and TC. SA's responses for the experiment in Chapter 8 were not recorded, due to equipment error, and TC was unavailable for testing for that experiment but participated in the two reported in this chapter. Participant details are shown in Table 9.3.

Table 9.3. Participant details

Measure		**G-SLI** N = 15	**LA1 controls** N = 12	**LA2 controls** N = 12	**LA3 controls** N = 12
Age	**Mean**	13;03	6;00	7;01	8;02
	Range	9;08 – 17;09	5;04 – 6;06	6;07 – 7;06	7;09 – 8;05
TROG	**Raw, mean**	12.47	14.53	16.17	17.17
	Raw, range	6 – 17	12 – 17	14 – 19	15 – 19
	z-score, mean	-1.70	0.59	0.59	0.47
BPVS	**Raw, mean**	75.07	68.80	76.92	91.00
	Raw, range	47 – 107	60 – 81	63 – 97	71 – 106
	z-score, mean	-1.84	0.55	0.35	0.49

In order to determine how the G-SLI group compares to the control groups on the two language measures, a series of independent samples t-tests was performed. For the TROG, the G-SLI group performs significantly worse than the LA1 controls, $t(25) = -2.156$, $p = 0.041$, the LA2 controls, $t(25) = -3.714$, $p = 0.001$ and the LA3 controls, $t(25) = -5.067$, $p < 0.001$. The LA1 controls provide the closest match to the G-SLI group in terms of grammar ability, but they still score significantly above the G-SLI group. This was unavoidable – it proved impossible to find typically developing children in the three schools I was testing at who scored as low as lowest-scoring G-SLI children. For the BPVS, the G-

SLI group do not score significantly differently to the LA1 group, t (25) = 1.248, p = 0.224, or the LA2 group, t (25) = -0.342, p = 0.735, but scored significantly worse than the LA3 group, t (25) = -2.854, p = 0.009. The LA2 group provides the best match in terms of vocabulary ability, and the LA3 group is included in order to show the developmental pattern amongst the typically developing children.

9.2.1.4. Predictions

I make no predictions as to how the G-SLI group will perform in comparison to their controls because previous studies have reported conflicting results in this regard (Section 9.1.2). However, I do predict that all groups will perform better on singular conditions than plural conditions, and that of the different plural conditions, all will perform worse on s-*ɪz*.

Given Oetting and Rice's anecdotal evidence that naming latencies are longer in SLI children, I predict that the G-SLI group will have longer naming latencies for the plural conditions than the control children.

9.2.1.5. Coding of responses

The response coding is devised on the basis that the first analysis uses responses whatever their phonological shape while the second and third analyses use responses whose phonological (metrical) shape matches that of the target response. Responses were coded with regards to both phonological shape and the accuracy of number marking, as follows. Examples of responses are given, with the target in brackets.:-

1) Target response, correct number marking, e.g. *goat* (*goat*), *bees* (*bees*)
2) Non-target response, correct phonological shape, correct number marking, e.g. *melon* (*lemon*), *houses* (*churches*)
3) Non-target response, incorrect phonological shape, correct number marking, e.g. *owl* (*parrot*), *chairs* (*benches*)
4) Target response, incorrect number marking, e.g. *carrots* (*carrot*), *table* (*tables*)
5) Non-target response, correct phonological shape, incorrect number marking, e.g. *wasp* (*bees*), *shoes* (*boot*)
6) Non-target response, incorrect phonological shape, incorrect number marking, e.g. *beads* (*necklace*), *pineapple* (*pears*)
7) No response

Since children sometimes changed their answer half-way through a response, making it impossible to determine whether correct number marking was used, the final response was accepted as the child's answer.

9.2.2. Results

3 different levels of analysis are reported:-

- Correct morphological marking for singular and plural targets, regardless of the phonological shape of the response. Using the numbers from the coding scheme in Section 9.2.1.4, this corresponds to 1 + 2 + 3 / 1 + 2 + 3 + 4 + 5 + 6.
- Correct morphological marking for plural targets according to the phonological shape of the target. Responses whose phonological shape does not match that of the target are not included. Using the numbers from the coding scheme in Section 9.2.1.4, this corresponds to 1 + 2 / 1 + 2 + 4 + 5.
- Reaction time analysis for singular and plural targets according to the phonological shape of the target, i.e. 1 + 2 / 1 + 2 + 4 + 5.

'No response' errors are not counted in the denominator. The number of 'no responses' was very low (G-SLI = 1.33%, LA1 = 2.78%, LA2 = 2.64%, LA3 = 1.11%). The rates of self-correction, for number marking and for lexical item, were likewise low (G-SLI = 2.00%, LA1 = 2.22%, LA2 = 2.64%, LA3 = 1.67%).

9.2.2.1. Analysis 1. Singular versus plural scores

In this analysis no attention is paid to the phonological form of the word supplied – the analysis just considers whether a word is supplied with the correct number marking. The results are set out in Table 9.4.

Table 9.4. % correct responses for singular and plural nouns

Condition		G-SLI	LA1	LA2	LA3
Singular	**Mean (SD)**	95.56 (5.86)	98.82 (1.95)	97.69 (3.32)	99.57 (1.00)
Plural	**Mean (SD)**	77.56 (18.05)	88.03 (8.27)	86.89 (9.90)	92.17 (8.92)

A 4 (Group: G-SLI, LA1, LA2, LA3) x 2 (Condition: singular, plural) ANOVA revealed significant main effects of group, $F(3, 47) = 5.503$, $p = 0.003$, and condition, $F(1, 47) = 41.357$, $p < 0.001$. The two-way interaction was not significant, $F(3, 47) = 1.368$, $p = 0.264$. Post hoc comparisons (Bonferroni-corrected) investigating the main effect of group found that the G-SLI group performs significantly worse than the the LA1 group, $p = 0.037$, and the LA3 group, $p = 0.002$, but no different to the LA2 group, $p = 0.115$. None of the pair-wise differences between the control groups was significant.

9.2.2.2. Analysis 2. Plural performance according to condition

This analysis considers correct responses to plurals within each condition. The counts are made on the basis of responses where the child provided the correct lexical item plus responses where an alternative lexical item was provided whose stem and inflected forms were of the same phonological shape as the target. The results are presented in Table 9.5.

Table 9.5. Correct production of plural forms according to condition

Condition		G-SLI	LA1	LA2	LA3
s-*z*	**Mean (SD)**	86.54 (16.42)	93.95 (5.77)	94.76 (11.25)	94.26 (12.63)
sw-*z*	**Mean (SD)**	81.20 (21.88)	88.21 (18.09)	85.34 (12.01)	93.28 (7.91)
s-*ɪz*	**Mean (SD)**	62.74 (28.20)	82.96 (18.76)	82.43 (14.96)	89.82 (11.28)

A 4 (Group: G-SLI, LA1, LA2, LA3) x 3 (Condition: s-*z*, sw-*z*, s-*ɪz*) ANOVA revealed main effects of group, $F(3, 47) = 3.822$, $p = 0.016$, and condition, $F(2, 94) = 12.489$, $p < 0.001$, but no two-way interaction. Post hoc comparisons (Bonferroni-corrected) revealed that the G-SLI group performed significantly worse than the LA3 controls, $p = 0.016$, but that no other pairwise group differences were significant. T-tests comparing differences in performance between pairs of conditions across participant groups found that performance was significantly better for the s-*z* compared to the sw-*z* condition, $t(50) = 2.312$, $p = 0.025$, the s-*z* compared to the s-*ɪz* condition, $t(50) = 4.349$, $p < 0.001$, and the sw-*z* compared to the s-*ɪz* condition, $t(50) = 3.389$, $p = 0.001$. In other words, the order of success is s-*z* > sw-*z* > s-*ɪz*.

9.2.2.3. Analysis 3. Naming latencies

Naming latencies were analysed just for G-SLI and LA1 group. This is because the LA1 controls were the closest to the G-SLI group in terms of overall performance on the plural conditions in Analysis 2, and were reasonably close to the G-SLI group in terms of both grammar and vocabulary abilities. Latencies were calculated for the s-ø, sw-ø, s-*z*, sw-*z* and s-*ɪz* conditions only. The 3σ- ø condition was omitted from the analysis because it has no phonological match amongst the plural conditions.

A large number of tokens had to be discarded from this analysis for several reasons. The most common reasons were that the child supplied incorrect number marking on the noun, or produced a target that had the correct number marking but which was of the wrong phonological shape, e.g. *volcano* for *mountain*, *handbags* for *purses* etc.

Other reasons included poor recording quality, and the child supplying an article or counting the objects e.g. *a lemon, three keys*. 48.67% of the data from the G-SLI group, including the entire data from two children (DA and DT) and 28.47% from the LA1 group, were discarded. Furthermore, as with all reaction time data, the distribution of timings is right-tailed (i.e. the distribution is skewed to the left). The majority of latencies were in the 1000-1500 ms range, with a mean of 1377.40 ms and a SD of 787.82 ms. Different experimenters use different criteria for deciding where the upper cut off should be, and there is no agreed convention. I decided to cut off at 4000ms (3.33 SD above the mean), which resulted in the inclusion of 99% of all the responses that were suitable for analysis. Naming latencies are shown in Table 9.6.

Table 9.6. Naming latencies (in milliseconds) for plural naming task

Condition		G-SLI	LA1
s-ø	**Mean (SD)**	1151.20 (259.80)	1201.56 (222.77)
sw-ø	**Mean (SD)**	1297.77 (233.94)	1313.83 (219.92)
s-*z*	**Mean (SD)**	1356.23 (299.58)	1433.57 (257.87)
sw-*z*	**Mean (SD)**	1353.69 (276.29)	1336.67 (262.42)
s-*ɪz*	**Mean (SD)**	1381.78 (318.08)	1422.56 (273.37)

A 2 (Group: G-SLI, LA1) x 5 (Condition: s-ø, sw-ø, s-*z*, sw-*z* and s-*ɪz*) ANOVA revealed a significant effect of condition, F (4, 88) = 4.603, p = 0.002, but not of group. The interaction between group and condition was not significant. A series of paired samples comparisons, designed to investigate further the effect of condition, revealed that naming latencies were significantly lower for the s-ø condition compared to all four other stimulus types, t (24) = -2.860, p = 0.009 for the sw-ø condition, t (24) = -3.936, p = 0.001 for the s-*z* condition, t (24) = -3.320, p = 0.003 for the sw-*z* condition, and t (23) = -3.626, p = 0.001 for the s-*ɪz* condition. No other pairwise comparisons reached significance. Therefore the naming latency is s-ø < sw-ø = s-*z* = sw-*z* = s-*ɪz*.

9.2. Present progressive elicitation task

9.3.1. Method

9.2.1.1. Verb stimuli

Three conditions, with 10 verbs in each condition, were selected for this task. The phonological characteristics of the verb stem in each condition are presented in Table 9.7. The full list of stimuli can be found in Appendix E.2.

Table 9.7. Verb conditions

Condition	Phonological characteristics of stem	Examples	Mean frequency*
s-*ing*	stressed monosyllabic	*hiding, dancing*	2.805
sw-*ing*	disyllabic, strong-weak stress	*whispering, balancing*	2.904
ws-*ing*	disyllabic, weak-strong stress	*applauding, returning*	2.935

* Frequency obtained from Francis and Kucera

Pictures of people of a variety of ages engaged in different activities were collected into a booklet, with one picture on each page. Pictures were taken mainly from the British Picture Vocabulary Scales (Dunn, Dunn, Whetton & Burley, 1997) with a few from the Clinical Evaluation of Language Fundamentals (Semel, Wiig & Secord, 1995).

9.3.1.2. Procedure

The present progressive form was elicited as follows:-

'We're going to look at some pictures of people doing things. I'm going to tell you what they like to do or what they need to do, and you're going to tell me what they are actually doing.

e.g. *This girl likes to dance to the music. Tell me what she is doing.*'

The expected response is, obviously, *She is dancing* (*to the music*).

Stimulus sentences are presented are presented in Appendix E.3. In each case the verb followed by either a direct object phrase or a prepositional phrase, which the child was encouraged to repeat. The child was also encouraged to use pronoun, as provided in the second sentence of the elicitation phrase, so that each potential auxiliary of the progressive construction was preceded by a monosyllabic, stressed item.

9.3.1.3. Participants

The same participants took part as in the plurality experiment, except for one child from the LA2 control group, who was not available for testing. See Section 9.2.1.2 for details of participants.

9.3.1.4. Predictions

The predictions for rates of *-ing* and auxiliary omission depend on the phonological shape of the stimulus. For *-ing*, production is predicted to be most successful for the s-*ing* condition. Omission is predicted to be greatest for the sw-*ing* condition because the affix is outside the trochaic foot. We might predict competition between *-ing* and stem-final weak syllable, because only one can occupy that place in the trochee. It is more difficult to make predictions for the ws-*ing* condition. On the one hand, *-ing* is part of a trochee, then suffix deletion is not predicted. On the other hand, the results of the non-word repetition task reported in Chapter 5 showed that the presence of an initial weak syllable affects syllabic and segmental accuracy elsewhere in the word. One prediction is that the presence of a stem-initial weak syllable will affect the realisation of inflection, by interacting with morphological complexity and resulting in omission. It is also possible that the initial weak syllable will be omitted, but this is not expected to occur at high levels, given the low rates of initial weak syllable deletion reported in Chapter 5. For auxiliary use, competition is predicted between the auxiliary and the initial weak syllable in the ws-*ing* condition (see Section 9.1.3), which will result in higher levels of auxiliary omission than for the s-*ing* and sw-*ing* conditions.

9.3.1.5. Coding of responses

- Correct, e.g. *mend* → *is mending*
- Correct, but omission of stem weak syllable, e.g. *hammer* → *is hamming* (instead of *hammering*), *explore* → *is ploring*
- Bare stem, e.g. *finish* → *is finish*
- Auxiliary omission, e.g. *relax* → *relaxing*
- Other, e.g. *direct* → *is directly* (i.e. incorrect suffix), *lead* → *is doing leading* (i.e. alternative construction), *race* → *were racing* (i.e. past progressive)

Minor segmental substitutions and metatheses are accepted as correct, as long as the auxiliary and *-ing* are supplied. For example, *is banceling* is accepted for *is balancing*, and *is applauring* for *is applauding*.

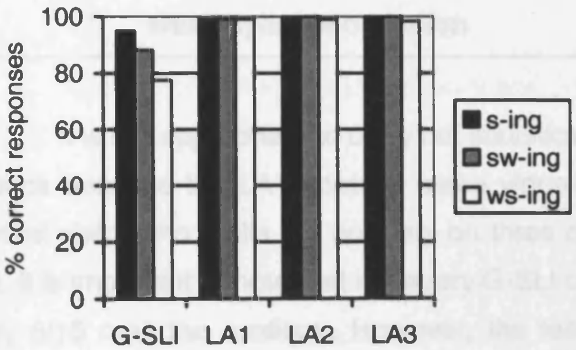
9.3.2. Results

First of all I analyse the correct scores from each group, according to verb stem-phonology, and then I perform an error analysis. The mean percentage of correct responses is shown in Table 9.8 and Figure 9.1. Correct responses are the sum of responses where the target is exactly correct, and those where the auxiliary and inflection are supplied, but where a stem weak syllable has been omitted.

Table 9.8. % correct responses

	Condition	G-SLI	LA1	LA2	LA3
s-*ing*	**Mean (SD)**	94.96 (14.90)	99.17 (2.89)	100.00 (0.00)	100.00 (0.00)
sw-*ing*	**Mean (SD)**	88.09 (24.86)	99.17 (2.89)	100.00 (0.00)	100.00 (0.00)
ws-*ing*	**Mean (SD)**	77.49 (32.86)	99.17 (2.89)	98.99 (3.35)	98.33 (3.89)

Figure 9.1. % correct responses



Because the LA2 and LA3 groups perform at ceiling, statistical analysis can only be performed on the G-SLI and LA1 groups. A 2 (Group: G-SLI, LA1) x 3 (Condition: s-*ing*, sw-*ing*, ws-*ing*) revealed a significant main effect of condition, F = 3.631, p = 0.034, but only a marginally significant effect of group, F = 3.783, p = 0.063. The two-way interaction was also significant, F = 3.631, p = 0.034. T-tests revealed that for the G-SLI group there were significant differences between the s-*ing* and ws-*ing* conditions, t (14) = 2.426, p = 0.029 and between the sw-*ing* and ws-*ing* conditions, t (14) = 2.700, p = 0.017. There were no significant differences between the s-*ing* and sw-*ing* conditions. T-tests are not possible for the LA1 group because their performance on all 3 conditions is identical.

Now I analyse the errors made on the progressive construction. The LA2 and LA3 groups perform at ceiling, so I consider just the G-SLI and LA1 groups. Recall that there are several different errors that children can make on the present progressive construction. *-Ing* omission is the morphological error, and this is predicted to vary as a function of the phonological complexity of the stem. However, two other errors are potentially influenced by phonology of stem – auxiliary omission (a syntactic error) and omission of a weak syllable in the stem. The results are presented in Table 9.9.

Table 9.9. Error types (% of total responses)

Condition	Error type	G-SLI	LA1
s-*ing*	***-ing* omission**	0.67	0.00
	auxiliary omission	4.70	0.83
sw-*ing*	***-ing* omission**	4.26	0.00
	auxiliary omission	2.84	0.83
	weak syllable omission	3.52	0.00
ws-*ing*	***-ing* omission**	9.70	0.00
	auxiliary omission	10.45	0.83
	weak syllable omission	3.73	4.27

It is not appropriate to carry out statistical comparisons between the G-SLI and LA1 groups because the LA1 controls make virtually no errors. Errors are made by just one control child, who omits the auxiliary on three occasions, for one verb from each stimulus set. It is important to note that not every G-SLI child makes errors – only 5/15 omit *-ing* and only 6/15 omit the auxiliary. However, the fact that several G-SLI children make these errors more than once is unexpected given their language age.

Non-parametric Wilcoxon signed ranks tests for the G-SLI group indicate that despite the relatively low proportion of errors, significant differences are found for *-ing* omission between the s-*ing* and ws-*ing* conditions, $Z = -2.023$, $p = 0.043$, and for auxiliary omission between the sw-*ing* and ws-*ing* conditionsi, $Z = -2.207$, $p = 0.027$. No other pairwise comparisons are significant. The picture is that ws-*ing* stimuli cause greater numbers of *-ing* and auxiliary omission errors.

9.4. Discussion: Metrical effects on inflectional morphology

9.4.1. Summary of results

9.4.1.1. Plural naming task

The performance of the G-SLI group relative to the control groups is not straightforward to interpret, given that the differences relative to the LA1 group are very nearly significant, but to the LA2 group are not. There is no developmental increase in performance between the LA1 and LA2 groups, and perhaps the customarily large standard deviations for the G-SLI group mean that significant differences do not emerge. However, the findings prevent us from claiming that the G-SLI group have a problem with plurality above and beyond what we would expect given their language age.

The phonological analysis confirms the findings of previous studies (Berko, 1958; Oetting & Rice, 1993) that /ɪz/ is harder than /z/, and this is true for both typically developing and G-SLI children. This study, however, further explores the difficulty with /ɪz/ by showing that nouns with two syllables are easier to inflect than those that take /ɪz/, and therefore that not all the difficulty with /ɪz/ nouns can be accounted for by difficulty of a two-syllable output.

The results of the naming latency analysis revealed that G-SLI children do not have longer naming latencies than the LA1 controls, and that the phonology of the stimulus affects both groups' performance in the same way. For both groups, naming latencies were shorter for the s-ø condition compared to the other four conditions. Importantly, even though nouns in the s-ɪz condition are inflected less frequently than those in the sw-z condition, when they are produced correctly they do not take longer to name.

9.4.1.2. Present progressive elicitation task

The difference in performance between the G-SLI group and the language-matched controls is much more clear-cut for the present progressive task: the controls are at ceiling and the G-SLI group is not. Effects of metricality are found for the G-SLI group but not for the controls. However, I suggest that this lack of metricality effect among the controls is likely to be due to me not having tested children at the period of development where metrical structure has an impact on present progressive formation, rather than the G-SLI children behaving differently with regards to phonology. The results indicate that G-SLI children do not have difficulty with *-ing* suffixation *per se*, but that omission occurs at significant levels if the stem has ws stress, i.e. is metrically complex. In terms of errors on the ws stem itself, the G-SLI group do not make more initial weak syllable deletions than the controls, confirming the findings in Chapter 5 that initial weak syllable deletion is not characteristic of G-SLI phonology.

9.4.2. Metrical effects on inflectional morphology

The two studies reported in this chapter were motivated by the need to further explore the effects of metrical structure on inflection. I discuss these first for plurality and then for present progressive formation.

The plural suffix /ɪz/ is added to stems ending in /s/ and /z/, which might be misanalysed as already being inflected for plurality. However, /ɪz/ is also added to /tʃ/, /dʒ/ and /ʃ/, which cannot be analysed as inflected. Performance on the 8 stems ending in /s/ and /z/, and the 4 ending in /tʃ/, is shown in Table 9.10.

Table 9.10. Mean bare stem responses on stems requiring /ɪz/, as a % of total responses for that stem form

Stem-final consonant	G-SLI	LA1	LA2	LA3
/tʃ/	31.37	15.38	13.04	4.65
/s/, /z/	36.19	19.72	21.95	12.50

Because this is a post hoc comparison, and because the stimuli were not designed for this purpose, I do not analyse the figures statistically. However, it is clear that bare stem forms appear almost as frequently for /tʃ/-final stems as for /s/ and /z/-final stems, at least for the G-SLI and LA1 groups. Therefore, if affix-checking does occur, it only plays a small part in explaining why bare stem forms occur so frequently for stems that require /ɪz/.

Interestingly, both the G-SLI and LA controls produce plurals that lack epenthesis e.g. **brushs* (Child KA). There are only a few such forms - 2 examples for G-SLI group and 3 for LA2 group. These errors have also been reported by Bernhardt and Stemberger (1998:642) and Berko (1958), but it is not known how often they occur in spontaneous speech. Most of the stimuli in this experiment end in /s/ or /z/, and instances of non-epenthesis could be underestimated because they are difficult to identify when the suffix is identical to the stem-final consonant. What these errors do show is that the children who make them are aware of the need to mark plurality, but resist adding /ɪz/, perhaps because they resist making a change to the metrical structure. The error is particularly interesting because sequences such as /ʃs/ are not phonotactically legal in English, in either monomorphemic or inflected words, and we know that G-SLI children at least are sensitive to the phonotactic patterns of word-final clusters (see the results of Chapters 3 and 4).

Another interesting observation is that the G-SLI group make no double plural marking errors (e.g. **bowses*), whereas they do make double tense marking errors on vowel-final stems (e.g. **cheweded*, see Chapter 8). No phonological reason springs to mind as to why children should double mark past tense but not plural forms. It may be due to differences in the demands of the two tasks – in the plural task, both singular and plural forms are elicited, whereas in the past tense task only past tense forms are elicited. Perhaps it is easier to adopt an explicit suffixation strategy in the past tense task. Alternatively, it may be that children's knowledge of the past tense is less secure than their knowledge of the plural.

The high levels of *-ing* suffixation are not predicted by my hypothesis that it is the change in metrical structure that accounts for problems using /ɪd/ and /ɪz/. There may exist a constraint against changing metrical structure that is dominated by a higher ranked morphological constraint whose role it is to ensure that *-ing* is realised. Because this amounts to no more than an ad hoc stipulation, I will not pursue the idea any further. Nevertheless, metrical complexity does affect the use of *-ing*, with more omission from ws-*ing* than from sw-*ing* verbs. The reason for this difference cannot be due to frequency differences between ws-*ing* and sw-*ing* verb-shapes. Kelly (1992) has shown that ws-*ing* verbs are more common than sw-*ing* verbs, presumably for the reason that sw-*ing* forms have a sequence of two unstressed syllables, which languages tend to avoid, whereas ws-*ing* verbs don't. The findings from the present progressive experiment are consistent with the CGC hypothesis, that impairments in syntax, morphology and phonology interact. Suppliance of the auxiliary is a syntactic process, as is the checking of aspectual features on *-ing*. Suffixation with *-ing* is a morphological process. When the phonology of the verb stem is kept metrically simple (i.e. one syllable), then the levels of auxiliary and *-ing* omission are very low, indicating that the syntax and morphology of present progressive formation is not impaired. However, when we introduce phonological complexity into the equation, the picture changes – auxiliary and *-ing* omission occur, particularly for the ws-*ing* stimuli. Is this omission purely a phonological phenomenon, or is it that morphological knowledge is less secure than in typically developing children, and therefore more likely to break down under stress from other aspects of language, e.g. phonology? We can tease these two possibilities apart because they predict different types of errors. If omission is purely a phonological phenomenon, *-ing* is predicted to be omitted from sw-*ing* stimuli more frequently than from ws-*ing* stimuli: in ws-*ing* stimuli the inflection is inside a trochaic template, w{s-*ing*}, whereas in {sw}-*ing* forms it is not. So phonology does *not* predict -*ing* omission in precisely the location where we get it most. This suggests that morphological knowledge is insecure, and the fact that there is negligible -*ing* omission in s-*ing* stimuli shows that it must be the presence of the initial weak syllable that is causing the difficulty. Why does the initial weak syllable have this effect on inflection?

We can draw parallels between the results for ws-*ing* verbs and the impact of initial weak syllables on non-word repetition (Chapter 5) – in both cases the presence of the weak syllable causes complexity elsewhere in the word to break down. This suggests that there is something about initial weak syllables that causes processing problems for any phonological material in the rest of the word, and those elements of phonological and morphological structure that are known to be difficult for children with SLI, e.g. consonant clusters, inflectional markers, break down.

Additional evidence that children find initial weak syllables problematic comes from one child in the G-SLI group, SA. SA uses */lə/* as the initial weak syllable in four of the ws-*ing* stimuli, producing *larresting, larranging, lappauding* and *labracing. Lappauding* might be considered an example of metathesis, but none of the other examples contain a */l/* in the stem. A more plausible explanation is that SA avoids producing word-initial weak syllables that lack an onset. De Lacy (2003) has argued that there is a constraint ONSETσ_1 that requires initial syllables to have onsets. This constraint must be so low-ranked, at least for initial unfooted syllables, as to be inactive in adult English, because vowel-initial words such as *applaud* and *arrange* exist (with no initial glottal stop, c.f. footed *apple* and *egg*, which do begin with a glottal stop). However, ONSETσ_1 could be higher ranked so as to be active in SA's grammar. This is a plausible explanation, but then we would expect *lembracing* rather than *labracing,* and would also predict that vowel-initial *emptying,* with strong-weak stress, would be produced as *lemptying* – but it isn't. An alternative explanation for the presence of */lə/* is that it is used as a dummy syllable when the child does not know the word and is therefore unsure of what segmental material goes there. Substitution of the segmental material by a dummy syllable has been reported for typically developing children (Gnanadesikan, 1995), but not to my knowledge in SLI. However, SA does not use */lə/* as a dummy syllable when the initial weak syllable begins with a consonant, which he might be predicted to do.

The work reported in this chapter has revealed that even for inflectional affixes that are relatively unimpaired in G-SLI children, such as the plural and progressive, phonology affects performance. In other words, complexity in one aspect of language can affect the realisation of complexity in another aspect, in this case leading to omission of -*s* and -*ing*, consistent with the CGC hypothesis that was detailed in Section 8.4.3. To test the CGC hypothesis further, in the next two chapters the effects of phonological complexity, and then inflectional complexity, will be tested on a little-studied area of morphology: derivational morphology.

Chapter 10. The impact of metrical structure on comparative and superlative formation

10.1. Introduction

10.1.1. Chapter outline

Few studies have been carried out on derivation in SLI, and this reflects a strong bias towards studies of inflectional morphology in the language acquisition field as a whole. This chapter reports on a task designed to elicit comparative and superlative adjectives that take *-er/-est*, in order to determine whether the omission of those derivational suffixes characterises G-SLI grammar in the same way that omission of the past tense suffix does. The metrical structure of the stimuli is varied in order to determine the effect of prosodic complexity on response accuracy.

In Section 10.1.2 I discuss the properties of derivational morphology versus those of inflection, and in Section 10.1.3 I discuss the properties of comparative/superlative formation, and previous studies of this phenomenon in SLI. In Section 10.2 I present the methodology employed in this study, and in Section 10.3 the results. In Section 10.4 I propose an Optimality-Theoretic account of the interactions between phonology and derivation. In Section 10.5.1 I summarise and discuss the results, and in Section 10.5.2 I discuss the data within the CGC framework, and begin to sketch an account of how the CGC can be extended into a developmental model. Part of the work in this chapter has been revised for *Language Acquisition* (Marshall & van der Lely, in prep.).

10.1.2. Derivational morphology

Linguistic theory distinguishes between (at least) two morphological processes – inflection and derivation. Although it has proved difficult to draw the distinction between inflection and derivation (see Section 1.2.2), Aronoff provides a useful definition: 'Inflection is the morphological realisation of syntax, while derivation is the morphological realisation of lexeme formation' (Aronoff, 1994:126). As a rule of thumb it is generally assumed that inflectional morphology is accessible to, and therefore manipulable by, syntax, because morphosyntactic properties are phrase-level properties to which syntactic relations such as agreement are sensitive (e.g. Anderson, 1982). Derivational morphology, on the other hand, is a lexical process. It serves to encode lexicosemantic relations within the lexicon, and is therefore inaccessible to the syntax. Some accounts hold that derivational forms are explicitly stored in the lexicon but regular inflected forms are not (e.g. the Split Morphology Hypothesis, Perlmutter, 1988)

A cognitive distinction between inflection and derivation has also been difficult to draw, and it is not clear whether inflected and derived forms are represented and processed in the same way or differently. Studies using repetition priming methodology have obtained a mixed pattern of results – some studies show larger effects of inflected primes than of derived primes, suggesting that inflected forms are more likely to be decomposed into stem and affix, but other studies show that inflected and derived forms have equivalent effects (see Raveh & Rueckl, 2000, for a review).

Studies of aphasic patients, however, provide evidence of a cognitive distinction between inflection and derivation. A number of patients have been documented with impairments in inflectional but not derivational morphology (see Badecker & Caramazza, 1998, for a review). A further way of determining whether a distinction between inflection and derivation exists is to compare their development. Children start to acquire inflection younger than derivation, at one and a half years for inflection as opposed to three years for derivation (Clark, 1998). However, many factors influence the rate of acquisition of inflectional affixes, including semantic complexity, allomorphy and the existence of irregularity. If inflection is acquired earlier than derivation then it is likely to be simpler than derivation in some respect. And yet, inflection requires attention to both lexical meaning and syntax, whereas derivation only requires attention to lexical meaning. On this basis, children who have difficulties with syntax, for example children with G-SLI, might acquire derivation more easily than inflection.

The proposed distinction between inflection (syntactic) and derivation (lexical), teamed with G-SLI children's proposed deficit in syntactic feature-checking and morphological suffixation, allows us to make predictions about the relative abilities of G-SLI subjects on derivational and inflectional morphology. We have seen that G-SLI children have difficulty with past tense formation, which is a type of inflection (see Chapters 6-8 of this thesis). This is unsurprising, as tense marking can be seen as part of their wider-ranging deficit in checking syntactic features via long-distance dependencies. In derivation, however, there is no such feature checking. Therefore this aspect of the impairment in G-SLI should not impact on derivational morphology. On the other hand, assuming that the Words and Rules model holds for productive derivation too, the deficit in morphological rule application will presumably have an impact on derivational affixation. In that case we would expect to find problems with derivation – but not to such an extent as in inflection, where syntactic feature-checking is also involved.

Could phonological complexity impact on derivational targets? Derivation does not build up verb-end complexity in the way that past tense inflection does, but it can change the metrical structure of the stem by adding a syllable (e.g. *happy* → *happiest*, *spot* →

spotty, sing → *singer, guitar* → *guitarist* etc.). Phonology's impact on derivation can be investigated just as it has been for inflection (Chapters 6-9).

10.1.3. Comparative and superlative formation

The experiment reported in this chapter elicits suffixed comparative and superlative adjectives. Although derivation is generally less productive than inflection, comparative and superlative suffixation is very productive: *-er* and *-est* attach to any gradable adjective of one syllable, or two syllables with strong-weak (sw) stress. There are a few, high frequency, irregulars, e.g. *good* → *better, bad* → *worse.* For non-gradable adjectives and those longer than two syllables, the phrasal route, e.g. *more real, most adorable,* is used. Presumably because of its high productivity, comparative and superlative formation has been classified as a type of inflection (e.g. Stump, 1998), although I follow others (e.g. Beard, 1998; Clahsen, Sonnenstuhl & Blevins, 2003) in considering it to be a derivational process. Graziano-King (1999) has claimed that there is no *-er* suffixation rule and that comparative forms are individually stored in the lexicon. However, given the high productivity and low age of acquisition of *-er,* I assume that comparative and superlative formation are rule-based processes, and note that this is not incompatible with the hypothesis that the products of these derivations may be stored in the lexicon (see Clahsen *et al.*, 2003).

Dalalakis (1994) has investigated comparative formation in children and adults from the KE family. She found that language-impaired members of the family had problems forming comparatives with *-er* (82% correct, compared to 93% for the controls) and *more* (21% correct, compared to 97% for the controls), but she provides no statistical analysis to show whether their performance on *-er* was significantly lower than that of the controls. There is no separate error analysis for adjectives taking *-er* and for those taking *more*, although Dalalakis reports a low proportion of bare stem errors overall, of just 15%. Therefore it is not clear that the language-impaired members have a deficit in *-er* marking that results in suffix omission.

Piggott and Kessler Robb (1999) asked members of the KE family to inflect a set of two-syllable nouns with the adjective-forming suffix *-al*, e.g. *margin* → *marginal*; *parent* → *parental*. The impaired family members managed to do this as easily as the unimpaired controls, but examination of the prosody of the derived forms showed quite dramatic differences between the two groups. Whereas, in the vast majority of cases, the unimpaired controls kept stress on first syllable of the derived word, impaired subjects produced a range of prosodic anomalies, including compound stress on the suffix (e.g.

/ri.dʒən#nʊ̩/ for *regional*), the insertion of an extra syllable before the suffix (e.g. */fæ.mɪn.ʔu̩.rʊ/* for *fragmental*) and stem truncation (e.g. */pɛ̩.sʊ/* for *personal*). In other words, although language-impaired subjects were able to derive words using *-al*, these derived words were phonologically anomalous. This suggests that the mechanism for derivational morphology is impaired in some way in these subjects. However, it should be noted that affected members of this family suffer from difficulties with prosody and articulation (caused by verbal dyspraxia) that other SLI subjects may not.

A third study relevant to the experiment reported here was carried out by Wauquier-Gravelines, Jakubowicz, Sauzet, Durand and Franc (1997) on agentive derivation in French, e.g. */ʃɑ̃t/* (*chante*, 'sing') → */ʃɑ̃tœʁ/* (*chanteur*, 'singer'). 8 SLI children aged 5;07-13;00 participated in their study. The authors were interested in whether the children could supply the correct consonant between stem and suffix in those verbs where the stem-final consonant is latent, i.e. is only heard when a vowel-initial suffix, in this case */œʁ/* is added, e.g. */dɔʁ/* (*dort*, 'sleep') → */dɔʁmœʁ/* (*dormeur*, 'sleeper'). Additionally, the number of syllables in the stimuli was varied (either one or two) in order to determine whether syllable number had an effect on retrieval of the latent consonant. Syllable number does indeed have an effect: on occasion, what should be a three-syllable target is reduced to two syllables, whereas two-syllable targets are produced with the correct number of syllables. Unfortunately no examples of outputs are given, and nor are error types clearly defined (there is no category of 'suffix omission', just 'suffix error', and there are no examples of outputs that would fall into this category). Therefore it is unclear whether the reduction of what should be a three-syllable output results from the suffix being omitted or the stem being truncated.

The experiment reported here is designed to investigate two issues. Firstly, do G-SLI children omit the comparative and superlative suffixes at the high levels that they omit the past tense suffix? Secondly, does increasing the number of syllables in the stem increase the rate of suffix omission for G-SLI children? The theoretical motives for investigating the first of these issues should be clear. The second issue requires more justification, however.

The output of derivation of a two-syllable stem (e.g. *happy*) is a three-syllable word of strong-weak-weak (sww) structure (e.g. *happier*). Not only is this longer than the output of derivation from a one-syllable stem, e.g. *sadder*, but it is structurally more complex in that the final weak syllable is unfooted, attached instead directly at the word level. In the two-syllable *sadder*, in contrast, the suffix can be incorporated into the trochaic foot with the stem.

Work by Marshall, Ebbels, Harris and van der Lely (2002), Gallon, Harris and van der Lely (2004) and work presented in Chapter 5 of this thesis reveals that in a non-word repetition task where non-words are systematically varied according to metrical complexity, G-SLI children repeat non-words with unfooted syllables less accurately than when all syllables are footed. However, I have also shown (see Chapter 5) that children rarely delete final unfooted syllables. At first glance, then, it seems unlikely that a three-syllable sww output should be subject to suffix omission, Wauquier-Gravelines *et al.*'s results notwithstanding. However, there is one fundamental difference between non-word repetition and the derivational processes being tested here – in the former the output is required to have the same metrical structure as the input, whereas in the latter the output is required to have a different metrical structure. Under the circumstances of derivation we might expect maximal word effects to emerge. Assuming in line with the standard acquisition literature that the minimal and maximal word in English is the trochaic foot (Allen & Hawkins, 1978; Demuth & Fee, 1995; Gerken, 1994) we predict pressure to produce a trochaic, i.e. sw, output. While this maximal word constraint would be dominated, and therefore inactive, when the input already contains an unfooted syllable (as in a non-word repetition task), it could become active in certain circumstances, e.g. during morphology. Such 'emergence of the unmarked' phenomena, whereby a marked structure that is tolerated in the language as a whole is not allowed to appear under particular circumstances, are well-documented in child and adult phonology (e.g. McCarthy & Prince, 1994).

If we predict maximal word effects on the output, then what should that output be? Kehoe (2000) reports that children's truncations preserve the stressed syllable and the word-final syllable. Pater and Paradis (1996) similarly report that typically developing children retain the first and third syllables in sww words. The examples they provide include *broccoli* → /**baki**/, *buffalo* → /**bʌfo**/, *sesame* → /**semi**/, *cinnamon* → /**sɪmɛn**/ and tricycle → /**twaikl**/. While the final rhyme (and word-final consonant, if one is present in the input) seems to always be preserved, the second onset appears to be chosen on the basis of sonority – the one with the lowest sonority is the one that is generally retained (but not always – see *sesame*).

None of Kehoe (2000) or Pater and Paradis' (1996) words is suffixed, so we cannot be sure whether to expect stem truncation or suffix omission in this experiment. From a purely phonological point of view we would predict that for targets such as *happier* and *happiest* the suffix, which is the third syllable, would be retained, in line with the morphemically simplex words above. This was also the case for the truncated *-al* forms in

Piggott and Kessler Robb's (1999) study. However, because the target is morphologically complex, we might expect the suffix to be dropped. Of course, in order to check that the suffix omission in this case is due to maximal word effects rather than to a deficit in morphology, performance on one- and two-syllable stimuli would need to be compared – if maximal word effects are in operation and impact on suffixation, then suffix omission will be higher for the two-syllable stimuli.

Comparative and superlative formation have not been previously studied in the G-SLI population, but some indication of the types of errors that these children are likely to make is given by their performance on two standardised language tests – the Word Structure subtest of the Clinical Evaluation of Language Fundamentals (CELF; Semel, Wiig & Secord, 1995) and the Grammatical Closure subtest of the Illinois Test of Psycholinguistic Abilities (ITPA; Kirk, McArthy & Kirk, 1968). These tests were carried out either as part of the initial screening process for inclusion into the G-SLI subgroup, or by speech and language therapists as part of the yearly battery of tests administered to children in special language units. Each test contains only one regular item that takes the -*er/-est* endings – the CELF contains *fast* and the ITPA contains *big*. Examples of errors on these items include:-

- suffix omission, e.g. *big* (instead of *bigger*), *fast* (instead of *fastest*).
- wrong suffix, e.g. *biggest* (instead of *bigger*)
- use of periphrastic construction when not required e.g. *most fast* (instead of *fastest*)
- double marking, e.g. *more faster* (instead of *faster*)
- semantic substitution, e.g. *enormous* (*biggest*), *largest* (instead of *biggest*)
- phonological error, e.g. *first* (instead of *fastest*).

These are therefore among the types of errors I predict will be made in this study.

10.2. Method

10.2.1. Adjectival Stimuli

The stimuli consist of 10 monosyllabic and 10 disyllabic (sw) adjectives. See Table 10.1 for the phonological characteristics of the two conditions and examples of stimuli. The full list of stimuli can be found in Appendix F.1.

Table 10.1. Adjectival stimuli

Condition	Phonological characteristics	Examples
One-syllable	1 syllable → 2 syllables	*short, dark, sad*
Two-syllable	2 syllables → 3 syllables	*happy, dirty, narrow*

Note that the stimuli in the two-syllable condition are a restricted set, all of which, with the exception of *narrow*, end in *y*. The two-syllable stimuli chosen here take the comparative and superlative suffixes, because the aim of the study is to investigate morphology rather than periphrastic constructions involving *more/most*. *-Y* is itself a derivational suffix (see Chapter 11), converting nouns into adjectives, e.g. *dirt* → *dirty*. Because it is possible that the presence of *-y* might cause difficulties with the selection of *-er/-est*, I took care to select equal numbers of adjectives (N = 5) which are decomposable into stem + *-y*, e.g. *dirty* and *curly*, and those that are not, e.g. *happy* and *heavy*, in order that the effect of this variable could be investigated.

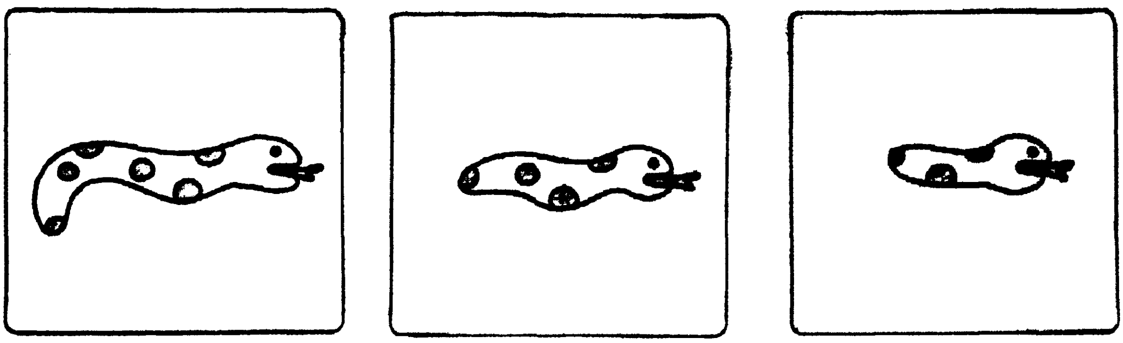
10.2.2. Procedure

The child is shown three pictures on a page (see Figure 10.1). The examiner points to each picture in turn and elicits the adjectives as follows:-

e.g. '*This snake is short, this snake is even ____(shorter) and this snake is the _____ (shortest)*'.

Two practice items were presented, for which corrections were provided if needed. The stimuli were randomised and one list created for all the participants.

Figure 10.1. Pictures for eliciting *shorter* and *shortest*



10.2.3. Participants

12 children with G-SLI, aged 9;10-16;08 (mean 12;01), participated in this experiment. Because of the wide range in language age demonstrated by the G-SLI group, and the discrepancy within individuals between grammatical and vocabulary abilities, each G-SLI child was individually matched to two typically developing children. 12 typically developing children were individually matched to G-SLI children on raw score (±1) obtained on the TROG (Bishop, 1983). A further 12 typically developing children individually matched to G-

SLI children on raw score (±3) obtained on the BPVS (Dunn, Dunn, Whetton & Burley, 1997). In order to get a picture of typical development, these control children were divided into two groups according to age. The LA1 control group are aged 4;06-6;11, with a mean age of 5;09, and the LA2 control group are aged 7;00-10;02, with a mean age of 8;09. Group participant details are shown in Table 10.2.

Table 10.2. Group participant details

Measure		**G-SLI** N = 12	**LA1 controls** N = 12	**LA2 controls** N = 12
Age	**Mean**	12;01	5;09	8;09
	Range	9;09 – 16;08	4;06 – 6;11	7;00 – 10;02
TROG	**Raw, mean**	12.33	10.25	15.92
	Raw, range	6 – 17	6 – 15	12 – 19
	z-score, mean	-1.74	-0.11	0.17
BPVS	**Raw, mean**	75.67	56.58	87.75
	Raw, range	47 – 104	44 – 70	69 – 102
	z-score, mean	-1.67	0.24	0.24

In order to determine how the G-SLI group compares to the control groups on the two language measures, a series of independent samples t-tests was performed. For the TROG, the G-SLI group does not perform significantly differently to the LA1 controls, $t(22) = 1.672$, $p = 0.109$, but performs significantly worse than the LA2 controls, $t(22) = -2.944$, $p = 0.008$. The LA1 controls therefore provide the closer match to the G-SLI group in terms of grammar ability. For the BPVS, the G-SLI group scores marginally worse than the LA2 group, $t(22) = -2.047$, $p = 0.053$, but performs significantly better than the LA1 group, $t(22) = 3.220$, $p = 0.004$. The LA2 group therefore provides a better match than the LA1 group in terms of vocabulary ability.

10.2.4. Predictions

The predictions for the G-SLI group, as framed within the CGC hypothesis, are two-fold. Firstly, given the difficulty with productive, rule-based morphology, I predict that G-SLI children will omit the derivational suffix. However, this omission will not be at rates as high as those for past tense omission, given that syntactic feature-checking is not required in derivational morphology. Secondly, given that G-SLI children show a deficit in at least some aspects of phonology, I predict they will show maximal word effects, but it is not clear whether these will come about through stem truncation or suffix omission. Typically developing children are not predicted to have difficulty with derivational suffixation. It is possible that the younger group may show maximal word effects, but as for the G-SLI group, it is not clear whether these effects will result in stem truncation or suffix omission.

10.2.5. Coding of the responses

Responses were coded as follows (the target is in brackets when it is not clear from the context what it should be):-

- Correct e.g. *silly → sillier, silly → silliest*
- Bare stem e.g. *silly → silly* (*sillier*)
- Stem truncation e.g. *tidy → tider, muddy → muddest*
- *More/most* + bare stem e.g. *funny → more funny*
- Others e.g. *curly → fluffiest, silly → silliness* (*silliest*), *heavy → more heaviest* (*heavier*)

A variety of identifiable error types come under the 'others' category, including semantic substitutions, selection of the wrong suffix and double marking through the use of *more/most* with a suffixed form. As the aim of this analysis is to focus on maximal word errors, I consider just three types of errors that are plausibly caused by maximal word constraints – bare stem, stem truncation, and *more/most* + bare stem – and lump the other errors in an 'others' category. Note that the use of *more/most* with a bare stem is not strictly speaking an error, given that the comparative or superlative is marked, but it is the inappropriate choice of construction for an adjective of this particular phonological shape.

10.3. Results

The proportion of correct responses for each group is presented in Table 10.3.

Table 10.3. % correct responses

Condition		**G-SLI**	**LA1**	**LA2**
Comp. 1 syll.	**Mean (SD)**	81.67 (33.26)	92.50 (10.55)	99.17 (2.89)
Comp. 2 syll.	**Mean (SD)**	61.67 (44.28)	77.50 (31.37)	94.17 (17.30)
Super. 1 syll.	**Mean (SD)**	89.17 (19.75)	84.17 (23.14)	97.50 (4.52)
Super. 2 syll.	**Mean (SD)**	66.67 (38.92)	80.83 (30.59)	100.00 (0.00)

To investigate whether the three groups show the same level of correct performance on the two different suffixes and on stimuli of different syllable number, a 3

(Group: G-SLI, LA1, LA2) x 2 (Suffix: *-er, -est*) x 2 (Syllable number: 1, 2) ANOVA was carried out. This revealed significant main effects of group, F (2, 33) = 3.462, p = 0.043 and syllable number, F (2, 33) = 8.593, p = 0.006. The main effect of suffix was not significant, and nor were any of the two-way interactions nor the three-way interaction. Post hoc comparisons (Bonferroni-corrected) were carried out to investigate the differences in performance between the groups. The G-SLI group performed significantly worse than the LA2 group, p = 0.040, but no differently to the LA1 group, p = 0.945. There was no significant difference in performance between the two control groups, p = 0.364. Because *-er* and *-est* behave the same way, and they are combined in the analysis that follows.

The finding that there are significantly fewer correct responses for two-syllable stimuli suggests that maximal word effects are present. In order to test this hypothesis, an error analysis was performed. Table 10.4 presents the error scores.

Table 10.4. Error types displaying maximal word effects, expressed as a % of total responses, comparatives and superlatives combined

Error type		**G-SLI**		**LA1**		**LA2**	
		1 syllable	**2 syllable**	**1 syllable**	**2 syllable**	**1 syllable**	**2 syllable**
Bare stem	**Mean**	2.08	0.83	0.83	8.33	0.00	0.42
	(SD)	(5.82)	(1.95)	(1.95)	(15.28)	(0.00)	(1.44)
Stem	**Mean**	n/a	22.08	n/a	2.50	n/a	0.00
truncation	**(SD)**		(33.74)		(3.371)		(0.0)
***More/most* +**	**Mean**	5.00	3.33	0.00	0.83	0.00	1.25
bare stem	**(SD)**	(9.29)	(6.15)	(0.00)	(1.95)	(0.00)	(4.33)

Because one of the error types, stem truncation, is not possible with one-syllable stimuli, it is not possible to carry out a group x error type x syllable number ANOVA. A 3 (Group: G-SLI, LA1, LA2) x 3 (Error type: bare stem, stem truncation, *more/most* + bare stem) ANOVA within just the two-syllable stimuli reveals a main effect of group, F (2, 33) = 4.655, p = 0.034, but no significant effect of error type. A significant group x error type interaction, F (2, 33) = 3.822, p = 0.007, indicates that the groups produce different patterns of errors on two-syllable stimuli. I now investigate each of these error types in turn.

For the bare stem errors on two-syllable stimuli, a one way ANOVA by group reveals only a marginally significant effect of group, F (2, 33) = 2.987, p = 0.064, indicating no real group differences in bare stem error production within two-syllable stimuli.

However, the question remains as to whether any of the groups produce bare stem forms as a way of avoiding a three-syllable output. A 3 (Group: G-SLI, LA1, LA2) x 2 (Syllable number: 1, 2) ANOVA reveals no significant main effects of group or syllable number, but a significant group x syllable number interaction, F (2, 33) = 3.413, p = 0.045, indicating that the groups produce different numbers of bare stem errors as a function of syllable number. To investigate this interaction further, t-tests within each subject group revealed that only the LA1 controls produced significantly more bare stem errors for two-syllable compared to one-syllable stimuli, t (12) = -1.827, p = 0.050 (1-tailed). This indicates that the LA1 controls, but no other groups, make bare stem errors in response to maximal word effects.

For stem truncation errors, a one way ANOVA reveals a significant group effect, F (2, 33) = 4.579, p = 0.018. Post hoc comparisons (Bonferroni-corrected) reveal that the G-SLI group make marginally more of these errors than the LA1 control group, p = 0.059, and significantly more than the LA2 group, p = 0.028. This indicates that only the G-SLI group respond to maximal word effects by producing stem truncation errors.

For *more/most* + bare stem errors, a one way ANOVA within the two-syllable stimuli showed no main effect of group. To check whether any of the groups are using this construction to avoid producing a three-syllable output, a 3 (Group: G-SLI, LA1, LA2) x 2 (Syllable number: 1, 2) ANOVA was carried out. This revealed no significant main effects of group or syllable number, and no significant group x syllable number interaction. The error analysis therefore reveals that G-SLI and LA1 children both show maximal word effects, but respond to these pressures in different ways: G-SLI children preferentially truncate the stem whereas LA1 children preferentially omit the suffix.

One obvious question at this point is whether children of any group are more likely to truncate a stem when that first syllable could stand as a semantically-related word on its own. We might predict that **mudder* would be produced more often than **heaver* because *mud* is a word in its own right whereas **heav* isn't. This factor was taken into account when the stimuli were chosen – five are 'decomposable' in that their truncated stem is a semantically-related word in its own right (*hairy, funny, curly, muddy, dirty*) and five are non-decomposable in that their truncated stem is not a word (*happy, silly, tidy, narrow, heavy*). Table 10.5 shows the number of stem-truncation errors apportioned to the decomposable and non-decomposable group.

Table 10.5. % stem truncation errors according to the decomposability of the stem

Stem type		G-SLI	LA1	LA2
decomposable	**Mean (SD)**	26.67 (38.46)	4.17 (6.69)	0.00 (0.00)
non-decomposable	**Mean (SD)**	16.67 (30.25)	0.83 (2.89)	0.00 (0.00)

The results indicate that G-SLI children truncate stems whether or not what remains of the stem is a word in its own right, but a t-test indicates that decomposable stems are more likely to be truncated than non-decomposable ones, $t\ (11) = 2.708$, $p = 0.020$. The LA children do not make enough stem truncations for statistical analysis to be reliable, but they too show a preference for truncating the stem when a real word results from that truncation. While the finding that G-SLI children are more likely to truncate on decomposable stems might suggest a morphological problem, the finding of such large numbers of truncations on non-decomposable stems indicates that stem truncations have a phonological cause.

10.4. Interactions between phonology and derivation: An Optimality Theoretic account

In this section I propose an OT account of the interaction between phonology and derivational morphology revealed in Section 10.3. There are three aspects of the G-SLI and LA1 data that need to be captured by such an account:-

- Both groups show maximal word effects.
- These maximal word effects are optional.
- The groups use different strategies: G-SLI – stem truncation; LA1 controls – suffix omission

Maximal word effects are obtained in this analysis by using the constraint MAX-Wdσσ. MAX-Wdσσ is defined as follows:-

MAX-Wdσσ - A word is maximally a bisyllabic trochee.

I recognise that this is a much-simplified cover term for a range of prosodic constraints, (e.g. De Lacy, 2003; McCarthy & Prince, 1995) but it is adequate for my purposes here. Two other constraints are required, in order to model the competition that arises between the stem-final syllable and the suffix. Which of these two syllables survives through to the output depends on relative ranking of the constraints, which I term REALISE-Suffix and MAX-Stem. They are defined as follows:-

REALISE-Suffix - Suffixes must be realised.

MAX-Stem - Don't delete material from the stem.

For G-SLI children it is more important to realise the comparative/superlative suffix, and therefore REALISE-Suffix is ranked higher. For typically developing children it is more important to maximise the stem, and MAX-Stem is ranked higher. These rankings are shown in Tableaus 10.1 and 10.2 below.

Tableau 10.1. G-SLI grammar

	/*hæpi-ə*/	REALISE-Suffix	MAX-Stem	MAX-Wdσσ
a. ☞	***hæpijə***			*
b. ☞	***hæpə***		*	
c.	***hæpi***	*!		

Tableau 10.2. LA1 grammar

	/*hæpi-ə*/	MAX-Stem	REALISE-Suffix	MAX-Wdσσ
a. ☞	***hæpijə***			*
b.	***hæpə***	*!		
c. ☞	***hæpi***		*	

The analysis is not as straightforward as it seems, however. One challenge is that for the LA1 children, error rates vary between the suffixes: out of a total of 16 suffix omissions, 14 involve *-er* and only 2 involve *-est*. This is not the case for truncation in the G-SLI group: there are 24 instances of truncation with *-er* and 29 instances of truncation with *-est*. The OT analysis therefore needs to capture the fact that it is easier for LA1 children to drop *-er* than to drop *-est*. One solution is to have two separate constraints, REALISE-*er* and REALISE-*est*, where REALISE-*est* is ranked higher than REALISE-*er*. Tableaus 10.3 and 10.4 show how this returns two candidates, /***hæpijə***/ and /***hæpi***/ as possible outputs for *happier*, but only one candidate, /***hæpijɪst***/, as a possible output for *happiest*. One justification for having a separate constraint for each suffix is that there are phonetic reasons why they might be acquired at different rates: /*ɪst*/ is more perceptually salient than /*ər*/ because it contains more segmental material.

Tableau 10.3. LA1 grammar, with separate constraints REALISE-*er* and REALISE-*est*

	/*hæpi-əl*/	MAX-Stem	REALISE-*est*	REALISE-*er*	MAX-Wdσσ
a. ☞	*hæpijə*				*
b.	*hæpə*	*!			
c. ☞	*hæpi*			*	

Tableau 10.4. LA1 grammar, with separate constraints REALISE-*er* and REALISE-*est*

	/*hæpi-ɪst*/	MAX-Stem	REALISE-*est*	REALISE-*er*	MAX-Wdσσ
a. ☞	*hæpijɪst*				*
b.	*hæpɪst*	*!			
c.	*hæpi*		*!		

For G-SLI children, REALISE-*er* and REALISE-*est* are presumably equally highly ranked, given that both are so rarely omitted. This ranking is shown in Tableau 10.5.

Tableau 10.5. G-SLI grammar, with separate constraints REALISE-*er* and REALISE-*est*

	/*hæpi-əl*/	REALISE-*est*	REALISE-*er*	MAX-Stem	MAX-Wdσσ
a. ☞	*hæpijə*				*
b. ☞	*hæpə*			*	
c.	*hæpi*		*!		

A second challenge is how to account for the fact that G-SLI children show maximal word effects in derivation but not in non-word repetition, e.g. *happier* → /*hæpəl*/ but *kɛtələ* → /*kɛtələl*/. This is a TETU ('The Emergence of the Unmarked') effect. TETU effects can be modelled using the framework of Comparative Markedness (McCarthy, 2003), which distinguishes between 'new' and 'old' versions of markedness constraints. An 'old' violation is one that is present in the fully faithful candidate, while a 'new' violation is one that is not present in the fully faithful candidate.

$_{O}$**MAX-Wdσσ** - A word is maximally a bisyllabic trochee. Applies to 'old' violations only.

$_{N}$**MAX-Wdσσ** - A word is maximally a bisyllabic trochee. Applies to 'new' violations only.

$_{N}$MAX-Wdσσ would be ranked higher than $_{O}$MAX-Wdσσ so that maximal word effects arise for derivational morphology but not for non-word repetition. Derivation produces 'new' violations, i.e. the auditory input to the child, e.g. *happy*, is a trochee and so does not violate MAX-Wdσσ, whereas the output, e.g. *happier*, does. In a repetition task, non-words contain 'old' violations, i.e. the auditory input to the child, e.g. ***kɛtələ***, is three syllables long and so already violates MAX-Wdσσ. These rankings are shown for derivation in Tableau 10.6 and for non-word repetition in Tableau 10.7.

Tableau 10.6. G-SLI grammar, with constraints $_{N}$MAX-Wdσσ and $_{O}$MAX-Wdσσ

	/***hæpi-ə***/	REALISE-*est*	REALISE-*er*	MAX-Stem	$_{N}$MAX-Wdσσ	$_{O}$MAX-Wdσσ
a. ☞	***hæpijə***				*	
b. ☞	***hæpə***			*		
c.	***hæpi***		*!			

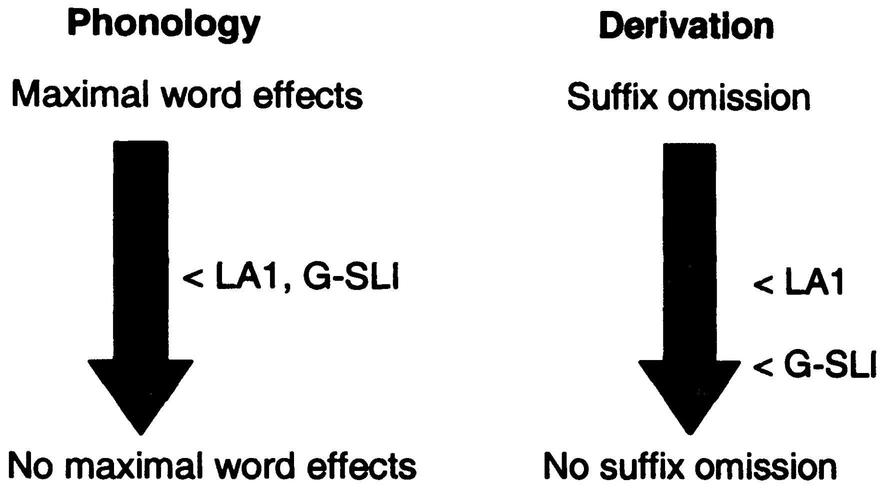
Tableau 10.7. G-SLI grammar, with constraints $_{N}$MAX-Wdσσ and $_{O}$MAX-Wdσσ

	/***kɛtələ***/	$_{N}$MAX-Wdσσ	MAX-C	MAX-V	$_{O}$MAX-Wdσσ
a. ☞	***kɛtələ***				*
b.	***kɛtə***		*	*	

There is a further explanation for the stem-truncation errors made by the G-SLI group, and that is that they are avoiding the vowel hiatus that would arise from having two adjacent nuclei, e.g. ***hæpiə*** (I thank an anonymous reviewer from *Language Acquisition* for pointing this out to me). Unfortunately my stimuli do not allow me to discount this possibility – I would need to check for the effects of vowel hiatus in a situation where maximal word effects would not be relevant, on one-syllable stimuli that end in a vowel, e.g. *blue*, *shy*. A maximal word account would predict correct forms such as *bluer*, *shiest*, whereas the hiatus account would predict truncated forms such as */blə/* and */ʃɪst/*. Even if the vowel hiatus account is the correct phonological explanation for the results reported in this chapter, the difference in morphological behaviour between the G-SLI and LA1 groups (retention versus deletion of the suffix for three-syllable targets) requires explanation.

Assuming that the maximal word effects analysis of the data is correct, what does this amount to in terms of phonological and morphological development? How can the different constraint rankings of REALISE-suffix relative to MAX- Wdσσ arise? I assume two separate developmental paths - one phonological, which takes the child from showing maximal word effects to showing no maximal word effects, and one morphological, which takes the child from omitting the derivational suffix to not omitting it. As both the G-SLI and LA1 groups show maximal word effects on occasion, they are presumably at the same stage along that particular pathway. In terms of suffix deletion however, the two groups are not at the same stage – suffix omission is more common for the LA1 than the G-SLI group. This means that in an environment when maximal word effects can impact on derivation, the two groups will produce different outputs. Hence in the case of the G-SLI group, 'abnormal' outputs can emerge from relative developmental timings that are different to the norm, and this is illustrated in Figure 10.2.

Figure 10.2. Developmental pathways for phonology and derivation



The idea that language development in SLI can be characterised by a delay in certain developmental pathways, such as tense marking, is not new (Rice, 2004; Rice, Wexler & Cleave, 1995). The new point that I make here is that abnormal productions can emerge from different relative developmental timings, contrary to the view that SLI children only produce forms that occur in typical development.

10.5. Discussion

10.5.1. Summary of results

Two findings emerge from this study. Firstly, G-SLI children rarely omit the comparative and superlative suffixes: omission rates are no higher than those of their language-matched peers, and are certainly nowhere near as high as omission rates for the past tense morpheme. Secondly, G-SLI children and typically developing children between the ages of 4;06 and 6;11 show maximal word effects. This means that an output that should be three syllables long is, on occasion, reduced to two syllables. G-SLI children and typically developing children achieve this reduction by different means – for G-SLI children, the preferred strategy is stem-truncation (e.g. /***hæpəl***/), whereas for typically developing children the preferred strategy is suffix omission (e.g. /***hæpil***/).

These two pieces of evidence indicate that omission of derivational suffixes is not a characteristic of G-SLI. Note that I am *not* claiming that derivation is unimpaired in G-SLI. The analysis presented in this study focused on suffix omissions and maximal word effects. However, the G-SLI children produced examples of double marking (e.g. *more sillier*) which are worthy of investigation, as they may indicate problems with the syntactic route of comparative/superlative formation. Indeed, in a group with syntactic difficulties such as the G-SLI group, errors with *more/most* might be predicted.

10.5.2 The status of derivational morphology

Children in both the G-SLI and LA1 groups both exhibit phonological pressure to reduce the output to the size of a trochaic foot, but which syllable is *actually* deleted is determined by how much pressure there is to retain the suffix – presumably for the G-SLI group the pressure is to keep the suffix, perhaps because of its semantic content. The LA1 group, on the other hand, is not under pressure to keep the suffix.

Do data from the non-word repetition test discussed in Chapter 5 shed light on which weak syllable of sww words G-SLI children are more likely to omit? Remember that weak syllable deletion is rare in these children. In fact, of the three children who show the greatest maximal word effects, QC and GS delete 0 out of 32 weak syllables in sww non-words, and OD deletes 5 out of 32. OD's omissions seem to be of the final syllable, e.g. ***fəkletələ*** $\rightarrow$ ***səkletə***, but this is not conclusive, because it is impossible to know whether the retained schwa is from final weak syllable or not, and whether /*t*/ is retained over /*l*/ for sonority reasons (see Section 10.1.3). In the derivational task OD retains the suffix, i.e. the last syllable, and deletes the middle syllable – he never omits the suffix. The fact that the control children show a different pattern of syllable deletion on that task – they delete the suffix – suggests that the choice of which syllable is to be deleted is down to morphological factors.

When G-SLI children truncate the stem in order to accommodate the derivational suffix within the minimum word, are they doing something outside the bounds of Universal Grammar, or does this phenomenon exist in adult language as well? A number of languages seem to impose conditions on maximal size of the word, but the phenomenon is little-researched. However, there is at least one example of stem-truncation in order to fit the suffix into an output that is limited in size by maximal word constraints. Yip (1992) discusses data from Anxiang, a dialect of Chinese. In Anxiang, diminutives are formed by reduplication and suffixation of */ər/* on the reduplicated syllable. The maximal word constraint in Anxiang is even more severe than that of English: it limits the size of the maximal word to a syllable. Replacing the rhyme of the reduplicated syllable by */ər/* leaves only the onset of the original word, and if the word has a high vowel this forms a glide in the onset, e.g. */mjan/* → */mjan mjər/* ('face').

The finding that *-er* is omitted more frequently by the LA1 controls than by the G-SLI group raises problems for any account of SLI that proposes difficulties in processing non-salient material (e.g. Leonard, 1989). I invoked a phonetic saliency account for why the LA1 children omitted *-er* more frequently than *-est* (see Section 10.4). However, if G-SLI children have difficulty processing non-salient suffixes, then why don't they omit *-er*? I argue that they don't omit it because they recognise its semantic importance. If this account is on the right lines, we can make an interesting comparison with the study of regular past tense inflection in Chapter 8. In that study, G-SLI children omitted */ɪd/* at the rate of 43.27%, and yet */ɪd/* has more phonetic material than */ə/*, and is therefore more salient. The striking difference in omission rates between */ɪd/* and */ə/* can be accounted for by the syntactic need for checking of tense features, an operation that is impaired in G-SLI grammar. An account such as the Surface Hypothesis, whereby SLI children omit suffixes of low phonetic salience, cannot account for their different behaviour with respect to */ɪd/* and */ə/*.

However, I also accounted for some of the omission of tense inflection being a result of an impaired suffixation rule. The findings in this chapter are problematic for the CGC hypothesis in this very respect. If we conceive of a morphological rule as taking a stem and adding a suffix to it, i.e. the operation stem + suffix, then surely a deficit in this rule should affect past tense and comparative/superlative formation alike. I argued for the existence of morphological rules, and their impairment in G-SLI, in Chapters 3 and 4, and yet from the findings reported in this chapter it appears that there is no rule impairment – - *er/-est* are not omitted. We expected to find lower rates of derivational suffix omission than past tense suffix omission because syntactic feature checking does not play a part in derivation, but the finding that derivational suffix omission is almost nil challenges the Words and Rules (WR) model (Pinker, 1999; Pinker & Ullman, 2002). If rule use is defective in inflection, then why is it not defective in derivation? The WR model predicts that (productive) derivation should be impacted as well (Pinker, personal communication, January 2004). The contrast between the deficit in inflectional suffixation and the lack of deficit in derivational suffixation strongly suggests that the WR model needs to be refined. It is too simplistic to have just one morphological rule of the form 'add a suffix to the stem'

– at the very least a distinction must be made between an inflectional rule and a derivational rule. Recall the effects of metrical complexity on the use of *-ing* and plural *-s* in Chapter 9. I argued that *-ing*, and perhaps *-s*, are not impaired *per se*, yet when the linguistic system is stressed through the addition of metrical complexity, those suffixes are prone to omission. The work presented in this chapter indicates that in the presence of metrical complexity inflection and derivation behave in different ways – inflectional suffixes are omitted and derivational suffixes are retained. This provides evidence for a cognitive distinction between the two.

Chapter 11. Further exploring derivational morphology – adjective formation from nouns

11.1. Introduction

11.1.1. Chapter outline

In Chapter 10 I presented data showing that children with G-SLI do not omit the comparative and superlative suffixes. This is in contrast to their characteristic omission of the past tense suffix, which was investigated in Chapters 3, 4, 6, 7 and 8. However, we also saw in Chapter 10 that increasing the phonological complexity of the adjectival stem leads on occasion to non-target responses, whereby a two-syllable stem is truncated under pressure to reduce the size of the output. These data can be accommodated within the CGC hypothesis, whereby complexity in components of the language faculty with which G-SLI children have difficulty can result in the production of non-target forms. This is despite the fact that derivational suffixation *per se* is not affected by increasing prosodic complexity.

In this chapter I probe derivation further in a study designed to elicit adjectives derived from nouns by the addition of the *-y* suffix (whose semantics mean 'having the characteristics that the noun refers to'). The aims of this study are to determine whether *-y* has similarly low levels of omission as *-er/-est*, and to investigate how the inflectional complexity of the stimulus affects the realisation of the target form. In Section 11.1.2 I discuss the derivation of adjectives and the theoretical motivation behind this study. In Section 11.2 I present the method and in Section 11.3 the results. In Section 11.4.1 I summarise the results, and in Section 11.4.2 I discuss how the CGC hypothesis can be extended to account for interactions between inflectional and derivational morphology. The work in this chapter has been revised for *Language Acquisition* (Marshall & van der Lely, in prep.).

11.1.2. The derivation of adjectives from nouns

The study presented in this chapter investigates two issues. Firstly, does the derivation of adjectives from nouns provide additional evidence for relative intactness of derivational suffixation in G-SLI? In other words, is the *-y* suffix used obligatorily in this context (as *-er* and *-est* are) or is it omitted (as the past tense suffix is?). Secondly, what is the impact on derivation of increasing the morphological complexity of the stimulus?

The investigation of the first issue should not require justification, given how few studies of derivation have been carried out on SLI children, but the second does require explanation. So far I have argued that children with G-SLI have difficulty with phonological complexity (Chapter 5), and shown that this difficulty impacts on derivational morphology in some children (Chapter 10). I have also argued that G-SLI children have difficulty with inflectional morphology, and discussed evidence from my own (Chapters 3 and 4) and other studies (e.g. van der Lely & Christian, 2000; van der Lely & Ullman, 2001) that they store inflected forms. So it is logically possible that this difficulty with inflection can impact on a different area of morphology: derivation. This possibility will be investigated through the derivation of adjectives from nouns by suffixation with *-y*, whose semantics mean 'having the characteristics of what the noun refers to' (e.g. 'a smoky room' has the characteristics of 'smoke').

The inflectional complexity of the noun stem is manipulated through the noun being either singular (e.g. *rain* → *rainy*) or plural (e.g. *spots* → *spotty*). Note that the adjective derived from *spots* loses the plural suffix when *-y* is added. This is an instance of the general rule that inflectional morphology does not occur within derivational morphology. Relevant here is the oft-made observation (e.g. Kiparsky, 1982) that while irregular plurals can appear inside compounds, e.g. *mice-eater*, regular plurals cannot, e.g. **rats-eater*. Kiparsky accounts for this phenomenon by ordering morphological rules into several levels, with the output of rules from later levels being unavailable for rules applying at earlier levels. He assumes that irregular plurals such as *mice* are formed in the lexicon at Level 1, whereas regular forms such as *rats* are created by an inflectional rule operating at Level 3. Compounding is assumed to take place at Level 2, i.e. after irregular plural formation but before regular plural formation. Given these assumptions, it follows that irregular plurals, but not regulars, can enter into compounding.

Pinker (1999) has proposed a psycholinguistically oriented theory based on level-ordering, in which morphologically irregular forms are stored in the mental lexicon whereas regular inflected forms are generated by rule. The rule governing the formation of inflected words applies after the rule governing compound formation. Although Kiparsky and Pinker's accounts differ in detail, they both assume that regular forms are not able to enter into compounding but that irregulars are. In support of their accounts, Gordon (1985) found that even children as young as 3 to 5 years old are sensitive to the different behaviour of regular and irregular nouns. When asked what they would call someone who eats rats, very few say **rats-eater*, but when asked what they would call someone who eats mice, they are happy to say *mice-eater*. This constraint against using regular plurals inside compounds poses a learnability problem because the constraint is evident at such a young age, while the frequency of compounds containing irregular plurals is near-zero. Children are therefore very unlikely to have derived their knowledge of this constraint from the input, and Gordon interprets this as indicating that the constraint is innate.

Van der Lely and Christian (2000) hypothesised, based on frequency effects found for regular past tense forms (van der Lely & Ullman, 2001), that G-SLI children store regularly inflected words in the lexicon. If this is the case, then G-SLI children should use regular plurals as well as irregulars inside compounds, given that both will be in the lexicon and available for compounding. They found that this was indeed the case for the majority of the G-SLI children that they tested: these children frequently produced forms such as **rats-eater*, whereas their language matched controls very rarely did. Van der Lely and Christian therefore concluded that G-SLI children store regularly inflected forms in the lexicon. Note, however, that it is not clear from this particular study whether G-SLI children really are storing regularly inflected forms or whether they lack knowledge of the rule-ordering process. The former explanation is more theoretically elegant, as there is independent evidence for the storage of morphologically regular forms: i.e. the frequency effects found for past tense regular verbs (van der Lely & Ullman, 2001). I shall therefore assume that the **rats-eater* type error can be accounted for by regular plural forms being stored in the lexicon, and thereby being available for compounding.

A serious criticism of experiment Gordon and van der Lely's compounding experiments, and the theory they are based on, is that regular plurals *do* occur inside a substantial number of compounds, e.g. *admissions committee, singles bar* and *drinks cabinet*. Space does not allow a full discussion of this criticism (see Alegre & Gordon, 1996; Haskell, MacDonald & Seidenberg, 2003; Pinker 1999; Ramscar, 2003). However, to my knowledge none of these authors has commented that although regular plurals do occur inside root compounds (whose right-most member is a noun, e.g. *singles bar, drinks cabinet*), they are very rare inside synthetic compounds (whose right-most word comprises a verb and affix, e.g. *rat-eater*). It may be the derivational process required for synthetic compounds that precludes the use of the regular plural, rather than the compounding itself.[4]

A further criticism (Stemberger, personal communication, November 2001) is that the design of Gordon and van der Lely and Christian's task, whereby the plural form is elicited before the compound, leads to priming effects. Presumably, in typically developing children only the irregular plural is primed, whereas for G-SLI children both irregular and regular plurals are primed. Stemberger's criticism may well be justified, but the difference

---

[4] Ramscar (personal communication, November 2003) counters with *arms dealer* as an example of a synthetic compound that contains a plural. However, he concedes that the plural is probably used to avoid ambiguity with singular *arm* meaning 'limb'.

in priming effects for regulars between typically developing and G-SLI children suggests that there is something fundamentally different about the way these two groups represent and process irregulars and regulars, with the typically developing children differentiating them and the G-SLI children treating them the same way.

Derivational morphology, like compounding, is proposed to take place in lexicon. The derivation of adjectives from nouns by the addition of *-y* can only apply to singular stems. A dress covered with spots may be described as *spotty* but not **spotsy*, even though something that is spotty must have more than one spot. Similarly a sea with lots of waves is *wavy* rather than **wavesy*, even though one wave does not make a wavy sea. However, if G-SLI children store plural forms such as *spots* and *waves* in their lexicon, then presumably they will have these plurals available for suffixation, and will produce forms such as **spotsy* and **wavesy*. On the other hand, typically developing children very rarely, if ever, will. A further prediction, based on our finding that the G-SLI group very rarely omitted the comparative and superlative suffixes, is that omissions of *-y* will be negligible.

As far as I am aware, the only example of a plural noun inside the *-y* affix is *gutsy*, meaning 'she/he's got guts' (there is also the slightly more risque *ballsy*). Presumably the plural is allowed here for semantic reasons – *gutsy* doesn't literally mean 'she/he's got lots of intestines' but rather 'she/he's got courage'. It is not possible to use 'she/he's got a gut' with the courage meaning – it would have to literally mean 'she/he's got an intestine'. So in *gutsy* the suffix has to be added to *guts* rather than to *gut* for semantic reasons.

As for irregular nouns, it appears that no irregular plurals occur before *-y*. We talk about a '*toothy* grin', not a '*teethy* grin', even though a grin needs more than one tooth in order to be toothy! A room infested with mice is *mousy* rather than *micy*, even though one mouse does not make an infestation. If indeed it is the case that irregular plurals do not occur inside *-y* either, then we have strong evidence that children are very unlikely to hear plurals of any kind inside *-y*. This means that the linguistic data are unambiguous: the child will receive no evidence that plurals of any kind can occur inside *-y*.

To my knowledge adjective-from-noun derivation has not been previously studied in either typically developing or SLI children. However, Clark reports a child (D) who at the age of 2;02 added *-y* to all the adjectives in his vocabulary, producing forms such as *darky* and *coldy*, before adding it to nouns a few weeks later to produce forms such as *crumby* and *cracky* (Clark, 2003). Hence it would appear that *-y* is acquired and used productively at a young age.

11.2. Method

11.2.1. Noun stimuli

There are 2 conditions, with 10 nouns in each condition. One condition consists of one-syllable singular nouns, the other of one-syllable plural nouns. The characteristics of these two conditions, and examples of stimuli, are presented in Table 11.1. The full list of stimuli can be found in Appendix G.1.

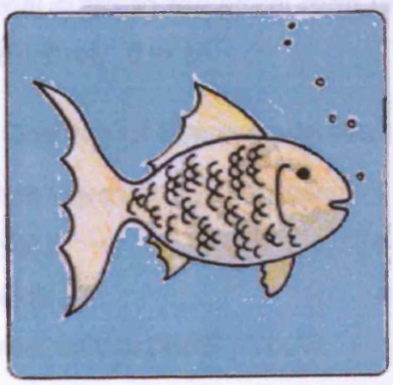
Table 11.1. Noun stimuli

Condition	Inflectional characteristics	Examples
singular	no plural inflection	*sun, mud, hair*
plural	plural inflection with /*s*/ or /*z*/	*spots, waves, rocks*

11.2.2. Procedure

The experiment is a simple elicitation task. The child is shown pictures, one on each page of a booklet. The lead in sentence, spoken by the examiner, is of the format, e.g. '*This fish has lots of scales*'. The next sentence, also spoken by the examiner, is designed to elicit the adjective, e.g. '*This fish is very* _______'.

Figure 11.1. Picture for eliciting '*This fish is very scaly*'



Two practice stimuli were presented and corrections given if necessary. The stimuli were randomised and one list order was created for all particpants.

11.2.3. Participants

The same participants participated in this study as in the comparative/superlative study (for details see Section 10.2.2). A summary of the three groups is given in Table 11.2.

Table 11.2. Group participant details

Measure		G-SLI N = 12	LA1 controls N = 12	LA2 controls N = 12
Age	**Mean**	12;01	5;09	8;09
	Range	9;09 – 16;08	4;06 – 6;11	7;00 – 10;02
TROG	**Raw, mean**	12.33	10.25	15.92
	Raw, range	6 – 17	6 – 15	12 – 19
	z-score, mean	-1.74	-0.11	0.17
BPVS	**Raw, mean**	75.67	56.58	87.75
	Raw, range	47 – 104	44 – 70	69 – 102
	z-score, mean	-1.67	0.24	0.24

11.2.4. Predictions

Based on the experiment reported in Chapter 10, where the G-SLI group very rarely omitted the comparative and superlative suffixes, I predict that the omission rates of *-y* will likewise be very low. I predict that typically developing children, on hearing a plural stimulus, will strip off the inflectional suffix before adding *-y* because they recognise that the regular plural is not a stem. In contrast, the G-SLI children may not strip off the plural suffix because if they have the plural form stored in their lexicon, it will be available for derivation. The prediction is therefore that G-SLI children will on occasion add the derivational suffix to the plural form of the noun, but that typically developing children will do so very rarely, if at all.

11.2.5. Coding of the responses

Responses were coded as follows:-

- Correct — e.g. *sun → sunny, spots → spotty*
- *-s* inside *-y* (applies only to plural stimuli) — e.g. *rocks → rocksy*
- Bare stem — e.g. *hair → hair, frills → frill, spots → spots*
- Others — e.g. *wool → fluffy, fur → furdy, rocks → rockily, holes →* no response

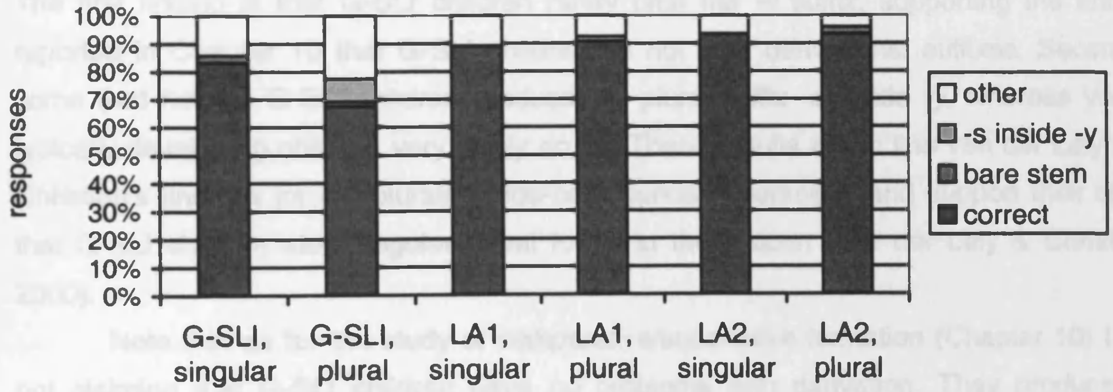
11.3. Results

All subjects understood the task and completed it quickly and fluently. The results are shown in Table 11.3.

Table 11.3. % responses for different response types

Response type		**G-SLI**		**LA1**		**LA2**	
		singular	**plural**	**singular**	**plural**	**singular**	**plural**
Correct	**Mean**	83.33	67.50	89.17	90.83	93.33	93.33
	(SD)	(16.14)	(29.58)	(14.43)	(15.64)	(7.79)	(8.88)
Bare stem	**Mean**	2.50	1.67	0.00	0.00	0.00	2.50
	(SD)	(6.22)	(3.89)	(0.00)	(0.00)	(0.00)	(8.66)
-s* inside *-y	**Mean**	n/a	8.33	n/a	1.67	n/a	0.00
	(SD)		(13.37)		(3.89)		(0.00)
Other	**Mean**	14.17	22.50	10.83	7.50	6.67	4.17
	(SD)	(13.11)	(29.27)	(14.43)	(12.15)	(7.79)	(5.15)

Figure 11.2. % responses for different response types



A 3 (Group: G-SLI, LA1, LA2) x 2 (Condition: singular, plural) ANOVA revealed main effects of group, $F(2, 33) = 4.529$, $p = 0.018$, and condition, $F(1, 33) = 4.233$, $p = 0.048$, and a significant group x condition interaction, $F(2, 33) = 5.903$, $p = 0.006$. The groups respond differently to plurality, and one way ANOVAS were used to unpack this interaction.

A one-way ANOVA within the singular condition revealed no significant effect of group, $F(2, 33) = 1.715$, $p = 0.196$. However, a one-way ANOVA within the plural condition did find a significant effect of group, $F(2,33) = 6.098$, $p = 0.006$. Post hoc comparisons (Bonferroni-corrected) reveal that the G-SLI group perform worse than both the LA1 controls, p = 0.022, and the LA2 controls, p = 0.010. Analysis of the simple effects was carried out using t-tests comparing performance on the singular and plural conditions within each group. Only the G-SLI showed a significant difference in performance on the two conditions, with the plural being harder than the singular condition, t (11) = 2.916, p = 0.014.

The mean percentage errors shown in Table 11.3 indicate that bare stem errors are very rarely produced by any of the groups. A 3 (Group: G-SLI, LA1, LA2) x 2 (Condition: singular, plural) ANOVA reveals no significant main effects of group or of condition on bare stem error production, and no significant interaction between group and condition. A second error type, the plural inside *-y* error is fairly obviously not expected to occur for nouns in the singular conditon. A one-way ANOVA reveals a significant main effect of group, F (2, 35) = 3.609, p = 0.038. Post hoc comparisons (Bonferroni-corrected) reveal a significant difference between the G-SLI group and the LA2 controls, p = 0.048, but not between the G-SLI and LA1 groups, p = 0.151.

11.4. Discussion

11.4.1. Summary of results

The first finding is that G-SLI children rarely omit the *-y* suffix, supporting the finding reported in Chapter 10 that G-SLI children do not omit derivational suffixes. Secondly, some (but not all) G-SLI children produce the plural suffix *-s* inside *-y*, whereas young typically-developing children very rarely do so. These results are in line van der Lely and Christian's findings for the plurals-inside-compounds experiment, and support their claim that G-SLI children store regular plural forms in the lexicon (van der Lely & Christian, 2000).

Note that as for the study of comparative/superlative formation (Chapter 10) I am not claiming that G-SLI children have *no* problems with derivation. They produced a substantial proportion of semantic substitutions (13.75%), more frequently, though not significantly so, than the control groups. This suggests that G-SLI children may have problems with lexical organisation, which are reflected in a higher incidence of substitutions. The strangest semantic substitutions certainly do come from the G-SLI children, with examples such as *scales* → **heavy* (where the picture stimulus shows scales on a fish, not weighing scales), *stars* → **heaven*, *rocks* → **dangerous*. What I *do* claim is that very little suffix omission occurs, and suffixation is remarkably robust to increases in stimulus complexity.

11.4.2. Extending the CGC hypothesis – interactions between inflectional and derivational morphology

I ended Chapter 10 by claiming that the different behaviour of inflectional and derivational suffixes in metrically complex environments provides evidence for a cognitive difference between inflectional and derivational suffixation. Evidence from this chapter shows again that derivational suffix omission is not characteristic of G-SLI, and we also have evidence that even under stress from inflectional complexity, *-y* is not lost.

The rates of *-s* inclusion are lower than in van der Lely and Christian's compounding experiment (8.33% in this experiment versus 35% in theirs). Not only are the rates lower, but a smaller proportion of children make this error (4/12 versus 14/16). These differences could arise because the outputs of adjectival derivation are likely to be lexicalised, whereas it is extremely unlikely that the compounds such as *rat-eater* and *mice-eater* are lexicalised. In my derivation experiment there are two ways of producing the correct item – *de novo* derivation and retrieval of the lexical item from the lexicon. In the compounding experiment the only available option is presumably *de novo* word formation. If we compare the frequencies of the adjectives in our experiment which are produced with *-s* inside, the majority have low frequencies. Higher frequency adjectives, such as *curly* and *cloudy*, are presumably more likely to be lexicalised, and therefore to be produced correctly.

Another reason for the lower-than-expected rates of *-s*-inside-*y* could be that the phonotactic sequences that result are often highly marked in terms of sonority (Heather Goad, personal communication, June 2004). For example, the */dz/* sequence that would arise from **cloudsy* and the */vz/* sequence from **wavesy* do not have the fall in sonority that English prefers in word-internal clusters (c.f. */ft/* in *after*, */lp/* in *pulpit*, */nd/* in *handle* etc.). Stop+fricative combinations such as */dz/* prefer to be voiceless, while fricative+fricative combinations such as */vz/* are just not well-formed. Note that 6 out of the 8 examples of the error have either no cluster (*starsy*) or */lz/*, where the sonority falls, while the other two have voiceless stop+fricative clusters. If this explanation is on the right tracks, it provides more evidence that G-SLI children are sensitive to cluster phonotactics (see Chapters 3 and 4).

Why do children make the **rocksy* type error? One possibility is that if the child has the plural form stored in the lexicon, this form will be available for derivational suffixation. This is the interpretation that van der Lely and Christian (2000) favoured for their compounding data. An alternative hypothesis is that the child does not analyse the plural input as being morphologically complex. Hay (2002) has developed a psycholinguistic model of affix selectional requirements, whereby constraints on the processing of morphological structure determine which suffixes will attach to which. For example, less parsable affixes cannot attach to more parsable affixes. My interpretation of Hay's proposal is that *-y* is less parsable than *-s* because *-s* is less likely to be analysed as being part of the stem, e.g. the final clusters of *curls*, *spots* and *clouds* are not found in monomorphemic words. Therefore *-s* (more parsable) cannot appear within *-y* (less parsable). Any child who is unable to parse the plural suffix will treat it as part of the stem, and produce plurals inside *-y*. Note that on some occasions where the G-SLI children include *-s*, the phonotactics are such that the inflection could be interpreted as part of the stem, e.g. *stars*, *rocks*. The higher incidence of *-s* inside *-y* errors in the G-SLI group may indicate that some have difficulty with morphological parsing.

When typically developing children produce plural-inside-*y* forms, do they make this error for the same reason as the G-SLI children or for a different reason? Are these forms acceptable in their grammar, or are they merely slips of the tongue? Grammaticality judgments of plural-inside-*y* forms in both G-SLI and typically developing children would enable us to test these two alternate hypotheses.

The CGC hypothesis requires a distinction between inflection and derivation, as reported in this chapter and the last. The work in both chapters has shown that in G-SLI derivational suffixation is remarkably resistant to omission, even when complexity in other areas of language is introduced.

Chapter 12. Conclusions

In this concluding chapter I summarise my research findings and the model that I have proposed to account for them – the Computational Grammatical Complexity hypothesis (Section 12.1). I outline some research questions that need to be answered if the model is to be developed further (Section 12.2), and I finish by suggesting some issues for linguistic and cognitive theory that arise from the work presented in this thesis (Section 12.3).

12.1. Summary

The aims of this thesis have been to investigate the phonological abilities of children with G-SLI and the various ways in which phonology impacts on morphology in both G-SLI and typical development. I began by demonstrating that at least part of the difficulty in producing regular past tense forms is a result of a morphological deficit. I showed that verb-end cluster phonotactics affect G-SLI children but not typically developing children, and I argued that this is a result of G-SLI children relying at least partially on the storage of past tense forms and/or their creation by analogy.

I then characterised the phonological deficit in G-SLI (Section 5.5.1) as being a deficit in the representation of complex phonological structure. In terms of syllable structure, the branching structure needed to represent consonant clusters is only optionally available. The impact of metrical complexity is harder to ascertain – certainly, children with G-SLI are able to produce unfooted syllables, but their presence at the left edge of the word causes errors at the syllabic and segmental levels. These errors are difficult to characterize, but I have made a start by investigating positional markedness effects on onset clusters (Section 5.4).

A crucial point is that not all children with G-SLI have a deficit in representing phonological complexity (Section 5.3.2), and so for these individuals, difficulties with past tense inflection cannot be caused by poor phonology. Indeed, some children make bizarre errors that can only be interpreted as revealing impaired morphosyntactic knowledge (Chapters 6 and 8). I therefore contend that a theory whereby phonological deficits cause the morphological deficit is too simplistic. Instead, I propose a model whereby deficits in syntax (as characterized by van der Lely, 1998), morphology (Chapters 3 and 4) and phonology (Chapter 5) impact on the realization of morphology. These deficits particularly hit regular past tense formation, which requires representations of syntactic, morphological and phonological complexity.

The impact of a phonological deficit can also be seen on areas of inflection that are not affected as badly as tense – plurality and the progressive (Chapter 9). In Chapters 10 and 11 I revealed an interesting dissociation between inflection and derivation. In the presence of phonological and syntactic complexity inflection is omitted, but in the presence of phonological and inflection complexity, derivation is still supplied.

I propose that the linguistic impairment in G-SLI lies in the representation of hierarchical complex structures in syntax, morphology and phonology. The complex structures that cause difficulty within each component include the following:-

- Syntax – non-local dependencies
- Morphology – concatenation of stem + suffix
- Phonology – consonant clusters and unfooted syllables

I term this model the Computational Gramatical Complexity (CGC) hypothesis. The CGC hypothesis can account for why tense is affected to such a great extent in the G-SLI subgroup, and across the SLI population more generally. In order to realise regular past tense forms, the child needs to have mastered complex structures in syntax (V to I movement of tense features), morphology (stem + *-ed*) and phonology (consonant clusters). Deficits in each of those three components will have an additive effect on past tense inflection, resulting in high levels of suffix omission.

12.2. Future research towards the development of the CGC model

The CGC hypothesis is in the early stages of development, and answers to the following questions will enable a fuller characterisation of the model:-

- Where exactly does the border lie between impairment and non-impairment in morphology? Past tense suffixation is undoubtedly impaired whereas derivational suffixation is not, yet the picture is less clear for plural *-s* and *-ing*. Finding what distinguishes impaired from unimpaired morphology will help us to better understand the locus of the morphological deficit.
- How do we account for the errors that are not reported for typically developing children, but that are made by at least some G-SLI children, e.g. 'am marrying a dancered?' (Section 8.3.2). First of all, do they really not occur in typically developing children? And if they do not occur in typical development, then how can the CGC model account for their presence in G-SLI?
- We need longitudinal studies in order to disentangle cause and effect. For example, auditory impairments are not characteristic of the G-SLI population (van der Lely, Rosen & Adlard, in press), but are they present at a younger age, and are they correlated with syntactic, morphological and/or phonological deficits? And why

is such a high co-occurrence of auditory-processing difficulties with SLI reported in the field as a whole?

- How does language knowledge interface with the performance systems, and what are the relations between representation and processing?
- Rather than working on group means in a population which has so much variability, we should be carrying out multiple case studies with full batteries of linguistic and cognitive tests, as advocated by Ramus (Ramus, 2003; Rosen, Dakin, Day, Castellote, White & Frith, 2003). This will enable us to determine what proportion of a particular group of SLI children has particular deficits, and how the severity of these deficits correlates with scores on experimental tasks. For this thesis I selected a group of children who have a syntactic/morphological deficit, but if I had instead selected children with phonological difficulties and then tested their syntax and morph, how would the picture look? Do all the combinations of dissociation between syntax, inflection, derivation and phonology exist?
- What are the semantic abilities of G-SLI children? As part of the grammatical computational system, semantics will need to be incorporated into a full account of the CGC model.

12.3. Issues for linguistic and cognitive theory –

The work presented in this thesis raises several issues for linguistics and cognitive science, and I outline what I consider to be the most important of these below:-

- What do syntactic, morphological and phonological complexity have in common? Are we dealing with an algorithm common to all components of grammar, which is impaired, or deficits in separate algorithms for complexity that have a high chance of co-occurring?
- When G-SLI children produce consonant clusters, do they have the same representation as in typically developing children? And when G-SLI children produce an inflected form, what is its morphological representation? Could it be that despite the target being correct, the neural mechanisms underlying representation and processing are atypical?
- What does it mean when a G-SLI child produces errors that are not made by typically developing children? Could these errors result from different rates of relative maturation in different areas of language (e.g. Section 10.4)?
- The variability in typically developing grammars needs to be acknowledged.

- How can we model interactions of constraints in OT? For example, what limits are there on the possible ways in which grammar can vary? Do we need to be thinking about constraints in terms of independent yet interacting pathways?
- When language breaks down in different populations, does it break down in similar ways, e.g. do language-impaired signers show effects of 'phonological' markedness for handshapes etc.?
- Ultimately our aim is to understand the full link between genes and phenotype, and in order to do this we need a much more detailed specification of the different phenotypes that display disordered language.

Finally, I hope to have convinced the reader that a close relationship between theory and language disorders benefits our understanding of both, and offers a fruitful avenue for research.

REFERENCES

Ackema, P. & Neeleman, A. (2000) Competition between syntax and morphology, in J. Grimshaw, G. Legendre & S. Vikner (eds) *Optimality theoretic syntax*. Cambridge, MA: MIT Press

Aguilar-Mediavilla, E.M., Sanz-Torrent, M. & Serra-Raventos, M. (2002) A comparative study of the phonology of pre-school children with specific language impairment (SLI), language delay (LD) and normal acquisition. *Clinical Linguistics and Phonetics*, 16: 573-596

Alegre, M.A. & Gordon, P. (1996) Red rats eater exposes recursion in children's word formation. *Cognition*, 60: 65-82

Alegre, M.A. & Gordon, P. (1999) Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, 40: 41-61

Allen, G.D. & Hawkins, S. (1978) The development of phonological rhythm. In A. Bell & J.B. Hooper (eds) *Syllables and segments*. New York, Elsevier-North Holland. pp. 173-185

American Psychiatric Association (1994) *Diagnostic and statistical manual of mental disorders IV*. Washington DC: American Psychiatric Association.

Anderson, S.R. (1992) *A-morphous morphology*. Cambridge: Cambridge University Press

Aram, D. & Nation, J. (1975) Patterns of language behaviour in children with developmental language disorders. *Journal of Speech and Hearing Research*, 18: 229-241

Aronoff, M. (1994) *Morphology by itself*, MIT Press, Cambridge, MA.

Baayen, H., Piepenbrock, R. & van Rijn, H. (1993) The CELEX Lexical Database, CD-ROM. University of Pennsylvania: Linguistic Data Consortium

Badecker, W. & Caramazza, A. (1998) Morphology and aphasia, in A. Spencer & A.M. Zwicky (eds) *The Handbook of Morphology*. Oxford: Blackwell

Baker, M. (1985) The mirror principle and morphosyntactic explanation. *Linguistic Inquiry*, 16: 373-416

Bauer, L. (1988) *Introducing linguistic morphology*. Edinburgh: Edinburgh University Press

Berko, J. (1958) The child's learning of English morphology. *Word*, 14: 150-177

Bernhardt, B.H. & Stemberger, J.P. (1998) *Handbook of phonological development. London*: Academic Press

Bird, H., Lambon Ralph, M.A., Seidenberg, M.S., McClelland, J.L. & Patterson, K. (2003) Deficits in phonology and past tense morphology: What's the connection? *Journal of Memory and Language*, 48: 502-526

Bishop, D.V.M. (1983). *The Test for Reception of Grammar.* Published by the author and available from Age and Cognitive Performance Research Centre, University of Manchester, M13 9PL

Bishop, D.V.M. (1994) Grammatical errors in specific language impairment: competence or performance limitations? *Applied Psycholinguistics*, 15: 507-551

Bishop, D.V.M. (1997) *Uncommon understanding: development and disorders of language comprehension.* Hove, UK: Psychology Press

Bishop, D.V.M., Bright, P., James, C., Bishop, S. & van der Lely. H.K.J. (2000) Grammatical SLI: a distinct subtype of developmental language impairment? *Applied Psycholinguistics*, 21: 159-181

Bishop, D.V.M., Chan, J., Adams, C., Hartley, J. & Weir, F (2000) Conversational responsiveness in specific language impairment: evidence of disproportionate pragmatic difficulties in a subset of children. *Development and Psychopathology*, 12: 177-199

Bishop, D.V.M. & Norbury, C.F. (2002) Exploring the borderlands of autistic disorder and specific language impairment: A study using standard diagnostic instruments. *Journal of Child Psychology and Psychiatry*, 43: 917-929

Bishop, D.V.M., North, T. & Donlan, C. (1996) Nonword repetition as a behavioural marker for inherited language impairment: evidence from a twin study. *Journal of Child Psychology and Psychiatry*, 37: 391-403

Blevins, J. (1995) The Syllable in Phonological Theory. In J. Goldsmith, (ed.), *Handbook of phonological theory.* London: Basil Blackwell. pp. 206-44.

Bloom, P. (2000) *How children learn the meanings of words.* Cambridge, MA: MIT Press.

Bortolini, U., Caselli, M.C. & Leonard, L.B. (1997) Grammatical deficits in Italian-speaking children with SLI. *Journal of Speech, Language and Hearing Research*, 40: 809-820

Bortolini, U. & Leonard, L.B. (2000) Phonology and children with specific language impairment: status of structural constraints in two languages. *Journal of Communication Disorders*, 33: 131-150

Bybee, J.L. & Slobin, D.I. (1982) Rules and schemas in the development and use of the English past tense. *Language*, 58: 265-289

Carter, A.K. & Gerken, L.A. (2003) Similarities in weak syllable omissions between children with SLI and normally developing language: a preliminary report. *Journal of Communication Disorders*, 36: 165-179

Chomsky, N. (1965) *Aspects of the theory of syntax.* Cambridge, MA: MIT Press

Chomsky, N. (1981) Principles and parameters in syntactic theory. In N. Hornstein & D. Lightfoot (eds) *Explanations in linguistics.* London: Longman. pp. 123-146

Chomsky, N. (1993) A minimalist program for linguistic theory. In K. Hale, & S.J. Keyser (eds) *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger.* Cambridge, MA: MIT Press. pp. 1-52

Chomsky, N. (1995) *The Minimalist Program.* Cambridge, MA: MIT Press

Clahsen, H. (1989) The grammatical characterisation of developmental dysphasia. *Linguistics,* 27: 897-920

Clahsen, H. & Almazan, M. (1998) Syntax and morphology in Williams Syndrome. *Cognition,* 68: 167 - 198

Clahsen, H., Bartke, S. & Göllner, S. (1997) Formal features in impaired grammars: a comparison of English and German SLI children. *Journal of Neurolinguistics,* 10: 151-171

Clahsen, H., Sonnenstuhl, I. and Blevins, J.P. (2003) Derivational morphology in the German mental lexicon: a dual mechanism account. In H. Baayen & R. Schreuder (eds) *Morphological structure in language processing.* Berlin: Mouton de Gruyter

Clahsen, H. & Temple, C. (2003) Words and rules in children with Williams Syndrome. In Y. Levy & J. Schaeffer (eds) *Language competence across populations: Towards a definition of specific language impairment.* Mahwah, NJ: Lawrence Erlbaum. pp. 323-352

Clark, E.V. (1998) Morphology in Language Acquisition, in A. Spencer and A.M. Zwicky, (eds) *The Handbook of Morphology.* Blackwell: Oxford. pp. 374-389

Clark, E.V. (2003) *First language acquisition.* Cambridge, UK: Cambridge University Press.

Cleave, P. & Rice, M. (1997) An examination of the morpheme BE in children with specific language impairment: the role of contractibility and grammatical form class. *Journal of Speech, Language and Hearing Research,* 40: 480-492

Coltheart, M. (1981) The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology,* 33: 497-505.

Coltheart, M. (1999) Modularity and cognition. *Trends in Cognitive Sciences,* 3: 115-120

Conti-Ramsden G, Botting N, Faragher B. J (2001) Psycholinguistic markers for specific language impairment. *Child Psychology and Psychiatry,* 42: 741-8

Cote, M.-H. (2002) *Consonant cluster phonotactics: A perceptual approach.* Doctoral dissertation, MIT. roa 548-1002

Cycowicz, Y.M., Friedman, D., Rothstein, M. & Snodgrass, J.G. (1997) Picture naming by young children: norms for name agreement, familiarity and visual complexity. *Journal of Experimental Child Psychology*, 65: 171 - 237

Dalalakis, J. (1994) English adjectival comparatives and familial language impairment. *McGill Working Papers in Linguistics*, 10: 50-66

Davies, L. (2001) *The nature of SLI: Optionality and principle conflict.* Unpublished doctoral thesis, University of London, England

De Bree, E. & van der Pas, B. (2003) Phonological complexity in children at risk for dyslexia and children with SLI. Poster presented at the Child Phonology conference, Vancouver, 1 - 4th July 2003

De Lacy, P. (2003) Maximal words and the Maori passive. In J. McCarthy (ed.) *Optimality theory in phonology: A reader.* Oxford: Blackwells. pp. 495-512

Demuth, K. (1996) The prosodic structure of early words. In J. Morgan & K. Demuth (eds) *Signal to syntax.* NJ: Lawrence Eribaum Associates. pp. 171-184

Demuth, K. & Fee, J. (1995) Minimal prosodic words in early phonological development. Manuscript, Brown University and Dalhousie University

Deruelle, C., Mancini, J., Livet, M.O., Casse-Perrot, C. & de Schonen, S. (1999) Configural and local processing of faces in children with Williams Syndrome. *Brain and Cognition*, 41: 276-98

De Villiers, J. (2003) Defining SLI: A linguistic perspective. In Y. Levy & J. Schaeffer (eds) *Language competence across population: Towards a definition of specific langage impairment.* Mahwah, NJ: Lawrence Erlbaum. pp. 425-447

Di Sciullo, A.-M. & Williams, E. (1987) *On the definition of word.* Cambridge, MA: MIT Press

Dockrell, J E., Messer, D. & George, R. (2001) Patterns of naming objects and actions in children with word finding difficulties. *Language and Cognitive Processes*, 16: 261-286

Dollaghan, C.A., Biber, M.E. & Campbell, T.F. (1995) Lexical influences on non-word repetition. *Applied Psycholinguistics*, 16: 211-222

Dollaghan, C. & Campbell, T. (1998) Nonword repetition and child language impairment. *Journal of Speech, Language and Hearing Research*, 41: 1136-1146

Dunn, L. M., Dunn, L. M., Whetton, C. & Burley, J. (1997) *British Picture Vocabulary Scales, 2nd edition.* Windsor: NFER Nelson

Dupoux, E., Kakeli, K., Hinose, Y., Pallier, C. & Mehler, J. (1999) Epenthetic vowels in Japanese: a perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25: 1568-1578

Ebbels, S. (2003) Phonology and argument structure in children with SLI: the influence of length versus syntactic complexity. Talk at City University, London, 23rd March 2003

Ebbels, S., van der Lely, H.K.J. & Dockrell, J. (2003) Phonological and morphological abilities in children with SLI: Is there a causal relationship? Talk presented at the 28th Annual Boston University Conference on Language Development, Boston, October 31 – November 2003

Echols, C.H. & Newport, E.L. (1992) The role of stress and position in determining first words. *Language Acquisition*, 2: 189-220

Edwards, J. & Lahey, M. (1998) Non-word repetitions of children with specific language impairment: Exploration of some explanations for their inadequacies. *Applied Psycholinguistics*, 19: 279-309

Elliott, C. D. (1996) *British Ability Scales, 2nd edition.* Windsor: NFER Nelson

Elman, J., Bates, E., Johnson, M., Karmiloff-smith A., Parisi, D. & Plunkett, K. (1996) *Rethinking innateness: a connectionist perspective on development.* Cambridge, MA: MIT Press

Field, A. (2000) *Discovering statistics using SPSS for Windows: Advanced techniques for the beginner.* London: Sage Publications

Fikkert, P. (1994) *On the acquisition of prosodic structure.* Holland Institute of Generative Linguistics.

Fodor, J. (1983) *The modularity of mind.* Cambridge, MA: MIT Press

Francis, W.N. and Kučera, H. (1982) *Frequency analysis of English usage: lexicon and grammar.* Boston: Houghton Mifflin.

Froud, K. (2001) *Linguistic theory and language pathology: Evidence for the morphology interface from a case of acquired language disorder.* Unpublished doctoral thesis, University of London

Froud, K. & van der Lely, H.K.J. (in prep.) Word learning in subgroups of specifically language impaired children: the independence of grammar and the lexicon. Manuscript, University College London

Gallon, N. (2002) *Prosodic complexity in specific language impairment.* Unpublished Masters thesis, Department of Phonetics and Linguistics, University College London

Gathercole, S.E. & Baddeley, A.D. (1990) Phonological memory deficits in language disordered children: is there a causal connection? *Journal of Memory and Language*, 29: 336-360

Gathercole S.E. & Baddeley, A.D. (1996) *Children's test of non-word repetition.* London: Psychological Corporation

Gavruseva, E. (2003) On the asymmetry in the development of copula and auxiliary *be* in child L2 English. Paper presented at the 28th Boston University Conference on Language Development, 31st October – 2nd November 2003

Gerken, L.A. (1991) The metrical basis for children's subjectless sentences. *Journal of Memory and Language*, 30: 431-51

Gerken, L.A. (1994) A metrical template account of children's weak syllable omissions from multisyllabic words. *Journal of Child Language*, 21: 565-584

German, D.J. (2000) *Test of Word Finding, 2nd Edition*. Austin, Texas: Pro-ed

Gleitman, L. & Wanner, E. (1982) *Language acquisition: the state of the art*. Cambridge, NY: Cambridge University Press

Gnanadesikan, A.E. (1995) Markedness and faithfulness constraints in child phonology. Rutgers Optimality Archive, roa67-0000.

Goad, H. (1998) Plurals in SLI: prosodic or morphological deficit. *Language Acquisition*, 7: 247-284

Goad, H., White, L. & Steele, J. (2003) Missing inflection in L2 acquisition: Defective syntax or L1-constrained prosodic representations? *Canadian Journal of Linguistics*, 1: 243-263

Goldsmith, J.A. (1976) *Autosegmental phonology*. Doctoral dissertation, MIT. Published 1979, New York: Garland

Gopnik, M. (1990) Feature-blindness: a case study. *Language Acquisition*, 1: 139-164

Gopnik, M. (1997) Language deficits and genetic factors. *Trends in Cognitive Sciences*, 1: 5-9

Gopnik, M. & Crago, M.B. (1991) Familial aggregation of a developmental language disorder. *Cognition*, 39: 1-50

Gopnik, M. & Goad, H. (1997) What underlies inflectional error patterns in genetic dysphasia? *Journal of Neurolinguistics*, 10: 109-137

Gordon, P. (1985) Level-ordering in lexical development. *Cognition*, 21: 73-93

Graziano-King, J (1999) *Acquisition of comparative forms in English*. Doctoral disssertation, CUNY Graduate Centre, New York.

Guy, G.R. (1991) Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change*, 3: 1-22

Halle, M. & Marantz, A. (1993) Distributed Morphology and the Pieces of Inflection. In K. Hale and S. J. Keyser (eds) *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, MA: MIT Press. pp. 111-176

Hamann, C., Penner, Z. & Lindner, K. (1998) German impaired grammar: The clause structure revisited. *Language Acquisition*, 7: 193-245

Harris, J. (1994) *English sound structure.* Oxford: Blackwell

Harris, J. & Gussman, E. (2002) Word-final onsets. *UCL Working Papers in Linguistics*, 14: 1-42

Haskell, T.R., MacDonald, M.C. & Seidenberg, M.S. (2003) Language learning and innateness: some implications of compounds research. *Cognitive Psychology*, 47: 119-163

Hay, J. (2002) From speech perception to morphology: Affix ordering revisited. *Language* 78, 527-555.

Hayes, B. (1995) *Metrical stress theory: principles and case studies.* Chicago: University of Chicago Press

Hsieh, L., Leonard, L.B., & Swanson, L.A. (1999). Some differences between English plural noun inflections and third singular verb inflections in the input: The contribution of frequency, sentence position, and duration. *Journal of Child Language*, 26, 531-543.

Ingram, D. (1981) Procedures for the phonological analysis of children's language. Baltimore, MD: University Park Press

Inkelas, S. & Rose, Y. (2003). Velar Fronting Revisited. In B. Beachley, A. Brown & F. Conlin (eds) *Proceedings of the 27th Annual Boston University Conference on Language Development.* Somerville, MA: Cascadilla Press. pp. 334-345

Ingham, R., Fletcher, P., Schelleter, C. & Sinkha, I. (1998) Resultative VPs and specific language impairment. *Language Acquisition*, 7: 87-112

Jakobowicz, C., Nash, L., Rigaut, C. & Gerard, C.-L. (1998) Determiners and clitic pronouns in French-speaking children with SLI. *Language Acquisition*, 7: 113-160

Jarrold, C. Baddeley, A. D. & Hewes, A.K. (2000) Verbal short-term memory deficits in Down Syndrome: A consequence of problems in rehearsal? *Journal of Child Psychology and Psychiatry*, 41: 233-244

Joanisse, M.F. (in press) Specific language impairments in children: phonology, semantics and the English past tense. *Current Directions in Psychological Science*

Joanisse, M.F. & Seidenberg, M.S. (1998) Specific language impairment: a deficit in grammar or processing? *Trends in Cognitive Sciences*, 2: 240-247

Joanisse, M.F. & Seidenberg, M.S. (1999) Impairments in verb morphology after brain injury: a connectionist model. *Proceedings of the National Academy of Sciences*, 96: 7592-7

Joanisse, M.F. & Seidenberg, M.S. (2003) Phonology and syntax in specific language impairment: evidence from a connectionist model. *Brain and Language*, 86: 40-56

Jusczyk, P. (1997) The *discovery of spoken language.* Cambridge, Massachusetts: MIT Press

Kail, R. (1994) A method for studying the generalised slowing hypothesis in children with Specific Language Impairment. *Journal of Speech and Hearing Research*, 37: 418-421

Karmiloff-Smith, A. (1998) Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences*, 2: 389-398

Karmiloff-Smith, A. & Thomas, M. (in press) What can developmental disorders tell us about the neurocomputational constraints that shape development? The case of Williams Syndrome. Special issue of *Development and Psychopathology*

Kehoe, M.M. (2000) Truncation without shape constraints: The later stages of prosodic acquisition. *Language Acquisition*, 8: 23-67

Kelly, M.H. (1992) Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review*, 2: 349-364

Kiparsky, P. (1982) From cyclic phonology to lexical phonology. In H. van der Hulst and N. Smith (Eds) The structure of phonological representations. Part 1. Dordrecht: Foris

Kirk, C. (2003) The acquisition of triconsonantal clusters in onsets and codas. Paper presented at the Child Phonology Conference, University of British Columbia, Vancouver, Canada, July 2nd- 4th 2003

Kirk, C. & Demuth, K. (2003) Onset/coda asymmetries in the acquisition of clusters. In . In B. Beachley, A. Brown & F. Conlin (eds) *Proceedings of the Boston University Conference on Language Development* 27. Somerville, MA: Cascadilla Press. pp. 437-448.

Kirk, C. & Seidl, A. (2003) Production and perception of unstressed initial syllables: implications for lexical representations. *Proceedings of the 28th Boston University Conference on Language Development*, 27. Somerville, MA: Cascadilla Press. pp. 318-327

Kirk, S., McArthy, J. & Kirk, W. (1968) *Illinois Test of Psycholinguistic Abilities*. University Press, Urbana, Illinois.

Kjelgaard, M. & Tager-Flusberg, H. (2001) An investigation of language impairments in autism: Implications for genetic subgroups. *Language and Cognitive Processes*, 16: 287-308

Lai, C.S.L., Fisher, S. E., Hurst, J.A., Vargha-Khadem, F. & Monaco, A.P. (2001) A fork-head domain gene is mutated in a severe speech and language disorder. *Nature*, 413: 519-523

Lasnik, H. (1995) Verbal morphology: syntactic structures meets the minimalist program. In H. Campos & P. Kempchinsky (eds) *Evolution and Revolution in Linguistic Theory.* Washington, DC: Georgetown University. pp. 251-275

Leonard, L. (1989) Language learnability and specific language impairment in children. *Applied Psycholinguistics*, 10: 179-202

Leonard, L. (1995) Functional categories in the grammars of children with specific language impairment. *Journal of Speech and Hearing Research*, 38: 1270-1283

Leonard, L. (1998) Children with specific language impairment. Cambridge, MA: MIT Press

Leonard, L., Bortolini, U., Casselli, M. C., McGregor, K. K. & Sabbadini, L. (1992) Morphological deficits in children with Specific Language Impairment: the status of features in the underlying grammar. *Language Acquisition*, 2: 151-179

Leonard, L., Deevy, P., Miller C., Charest, M., Kurtz, R. & Rauf, L. (2003) The use of grammatical morphemes reflecting aspect and modality by children with specific language impairment. *Journal of Child Language*, 30: 769-795

Leonard, L., Eyer, J.A., Bedore, L.M. & Grela, B.G. (1997) Three accounts of the grammatical morpheme difficulties of English-speaking children with Specific Language Impairment. *Journal of Speech, Language and Hearing Research*, 40: 741 - 753

Leonard, L., Sabbadini, L., Volterra, V. & Leonard, J. (1987) Specific Language Impaired Children: a cross linguistic study. *Brain and Language*, 32: 233-252

Levelt, C.C., Schiller, N.O. & Levelt, W.J. (2000) The acquisition of syllable types. *Language Acquisition*, 8: 237-264

Levy, Y. & Kave, G. (1999) Language breakdown and linguistic theory: A tutorial overview. *Lingua*, 107: 95-143

Lieber, R. (1992) *Deconstructing morphology*. Chicago: University of Chicago Press

MacWhinney, B. (1978) The acquisition of morphophonology. *Monographs of the Society for Research in Child Development, 43*

Marchman, V.A. (1997) Children's productivity in the English past tense: the role of frequency, phonology and neighbourhood structure. *Cognitive Science*, 21: 283-304

Marchman, V.A., Wulfeck, B. & Weismer, S.E. (1999) Morphological productivity in children with normal language and SLI: a study of the English past tense. *Journal of Speech, Language and Hearing Research*, 42: 206-219

Marcus, G.F. & Fisher, S.E. (2003) *FoxP2* in focus: What can genes tell us about speech and language? *Trends in Cognitive Sciences*, 7: 257-262

Marcus, G.F., Pinker, S., Ullman, M., Hollander, M., Rosen, T.J. & Xu, F. (1992) Overregularisation in language acquisition. *Monographs of the Society for Research in Child Development*, 57: 1-182

Marshall, C.R. & Chiat, S. (2003) A foot domain account of prosodically-conditioned substitutions. *Clinical Linguistics and Phonetics*, 17: 645-657

Marshall, C.R., Ebbels, S., Harris, J. & van der Lely, H.K.J. (2002) Investigating the impact of prosodic complexity on the speech of children with Specific Language Impairment. In R. Vermeulen & A. Neeleman, (eds) *UCL Working Papers in Linguistics*, 14: 43-68

Marshall, C.R., Harris, J. & van der Lely, H.K.J. (2003) The nature of phonological representations in children with Grammatical-Specific Language Impairment (G-SLI). In D. Hall, T. Markopoulos, A. Salamoura & S. Skoufaki (eds) *Proceedings of the University of Cambridge First Postgraduate Conference in Language Research*: 1: 511-517

Marshall, C.R. & van der Lely, H.K.J. (in prep.) Derivational morphology in children with Grammatical-SLI. *Language Acquisition:* Revisions submitted.

Marslen-Wilson, W.D. & Tyler, L.K. (1997) Dissociating types of mental computation. *Nature*, 387: 592-594

Marslen-Wilson, W.D. & Tyler, L.K. (1998) Rules, representations and the English past tense. In *Trends in Cognitive Sciences*, 2: 428-435

Matthews, J. (1994) Editor. Linguistic aspects of familial language impairment. *McGill Working Papers in Linguistics*, 10

McCarthy, J.J. (2003) Comparative markedness. *Theoretical Linguistics*, 29: 1-51

McCarthy, J.J. (2004) Optimal paradigms. In L. Downing, T.A. Hall & R. Raffelsiefen, (eds), *Paradigms in Phonological Theory*. Oxford: Oxford University Press

McCarthy, J.J. & Prince, A. (1995) Faithfulness and Reduplicative Identity. In J.I Beckman, S. Urbanczyk and L.W. Dickey (eds) *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*. pp. 249-384

McClelland, J.L. & Patterson, K. (2002) Rules or connections in past-tense inflections: what does the evidence rule out? *Trends in Cognitive Sciences*, 6: 465-472

McDaniel, D. & Smith Cairns, H. (1996) Eliciting judgements of grammaticality and reference. In D. McDaniel, C. McKee & H. Smith Cairns (eds) *Methods for assessing children's syntax*. Cambridge, MA: MIT Press. pp. 233-254

Mervis, C.B., Morris, C.A., Bertrand, J. & Robinson, B.F. (1999) Williams Syndrome: Findings from an integrated program of research. In H. Tager-Flusberg (ed.) *Neurodevelopmental disorders*. Cambridge, MA: MIT Press. pp. 65-110

Mervis, C.B. & Robinson, B.F. (2003) Methodological issues in cross-group comparisons of language and cognitive development. In Y. Levy & J. Schaefer (eds) *Language competence across populations: Towards a definition of Specific Language Impairment*. Mahwah, NJ: Lawrence Erlbaum. pp. 233-258

Miller, C.A. & Leonard, L.B. (1998) Deficits in finite verb morphology: some assumptions in recent accounts of specific language impairment. *Journal of Speech Language and Hearing Research*, 41: 701-707

Montgomery, J.W. & Leonard, L.B. (1998) Real-time inflectional processing by children with specific language impairment: Effects of phonetic substance. *Journal of Speech, Language and Hearing Research*, 41: 1432-1443

Nimmo, L.M. & Roodenrys, S. (2002) Syllable frequency effects on phonological short term memory tasks. *Applied Psycholinguistics*, 234: 643-659

Norbury, C.F. (2003) SLI and autism: Category, continuum or conundrum? Talk presented at the Institute of Child Health, University College London, 19th June 2003

Oetting, J.B. & Horohov, J.E. (1997) Past-tense marking by children with and without specific language impairment. *Journal of Speech, Language and Hearing Research*, 40: 62-74

Oetting, J. B. & Rice, M. L. (1993) Plural acquisition in children with Specific Language Impairment. *Journal of Speech and Hearing Research*, 36: 1236-1248

Orsolini, M., Sechi, E., Maronato, C, Bonvino, C. & Corcelli, A. (2001) Nature of phonological delay in children with specific language impairment. *International Journal of Language and Communication Disorders*, 36: 63-90

Peiris, D. (2000) *The influences of prosodic complexity on segmental production in SLI: A single case study*. Unpublished Masters thesis, Department of Human Communication Science, University College London

Pater, J. and Paradis, J. (1996) Truncation without templates in child phonology, in A. Stringfellow, D. Cahama-amitay, E. Hughes & A. Zukowski (eds) *Proceedings of the 20th Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press. pp. 540-551

Perlmutter, D. (1988) The split-morphology hypothesis: Evidence form Yiddish. In M. Hammond & M. Noonan (eds) *Theoretical morphology: Approaches in modern linguistics*. San Diego, California: Academic Press. pp. 79-100

Piggott, G. L. (1999) At the right edge of words. *The Linguistic Review*, 16: 143-185

Piggott, G. L. & Kessler Robb, M. (1999) Prosodic features of familial language impairment: Constraints on stress assignment. *Folia Phoniatrica et Logopaedica*, 51: 55-69

Pinker, S. (1984) *Language learnability and language development*. Cambridge, MA: Harvard University Press

Pinker, S. (1991) Rules of Language. *Science*, 253: 530-535

Pinker, S. (1994) *The Language Instinct*. New York: William Morrow

Pinker, S. (1999) *Words and rules: The ingredients of language.* London: Weidenfeld & Nicolson

Pinker, S. & Prince, A. (1988) On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28: 73-194

Pinker, S. & Ullman, M. (2002) The past and future of the past tense. *Trends in Cognitive Sciences*, 6: 456-463

Plunkett, K. & Juola, P. (1999) A connectionist model of English past tense and plural morphology. *Cognitive Science*, 23: 463-490

Plunkett, K. & Marchman, V.A. (1991) U-shaped learning and frequency effects in a multi-layered perceptron: implications for child language acquisition. *Cognition*, 38: 43-102

Prince, A. & Smolensky, P. (1993) Optimality Theory: Constraint interaction in generative grammar. *Rutgers University Centre for Cognitive Science Technical Report 2*

Radford, A. (1997) *Syntactic theory and the structure of English: A minimalist approach.* Cambridge, UK: Cambridge University Press

Ramscar, M. (2003) When and why might children say mice-eater? Paper presented at the 28th Boston University Conference on Language Development, 31st October – 2nd November 2003

Ramus, F. (2003) Developmental dyslexia: specific phonological deficit or general sensorimotor dysfunction. *Current Opinion in Neurobiology*, 13: 212-218

Ramus, F., Rosen, S., Dakin, S.C., Day, B.L., Castellote, J.M., White, S. & Frith, U. (2003) Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain*, 126: 841-865

Rapin, I. & Allen, D. (1983) Developmental language disorders: Nosologic considerations. In U. Kirk (ed.) *Neuropsychology of language, reading and spelling* New York: Academic Press. pp.155-184

Raveh, M. & Rueckl, J.G. (2000) Equivalent effects of inflected and derived primes: Long-term morphological priming in fragment completion and lexical decision. *Journal of Memory and Language*, 42: 103-119

Raven, J.C. (1998) *Raven's progressive matrices.* Oxford: Oxford Psychologists Press, Ltd.

Rice, M.L. (2003) A unified model of specific and general language delay: Grammatical tense as a clinical marker of unexpected variation. . In Y. Levy & J. Schaeffer (eds) *Language competence across populations: Towards a definition of specific language impairment.* Mahwah, NJ: Lawrence Erlbaum. pp. 63-94

Rice, M.L. (2004) Language growth of children with SLI and unaffected children: Timing mechanisms and linguistic distinctions. Plenary talk at the BU Conference on Language

Development. In A. Brugos, L. Micciulla, & C. Smith (eds.) *Proceedings of the 28th Annual Boston University Conference on Language Development,* Somerville, MA: Cascadilla Press. pp. 28-49

Rice, M.L., Tomblin, J.B., Hoffman, L., Richman W.A. & Marquis, J. (in press) Grammatical Tense Deficits in Children with SLI and Nonspecific Language Impairment: Relationships with Nonverbal IQ over Time. *Journal of Speech, Language and Hearing Research*

Rice, M.L. & Wexler, K. (1996) Towards tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research,* 39: 1239-1257

Rice, M. L., Wexler, K. & Cleave, P. L. (1995) Specific Language Impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research,* 38: 850-863

Rice, M.L., Wexler, K., Marquis, J. & Hershberger, S. (2000) Acquisition of irregular past tense in children with Specific Language Impairment. *Journal of Speech, Language and Hearing Research,* 43: 1126-1145

Rice, M.L., Wexler, K. & Redmond, S.M. (1999) Grammaticality judgements of an extended optional infinitive grammar: evidence from English-speaking children with specific language impairment. *Journal of Speech, Language and Hearing Research,* 42: 943-961

Roark, B. & Demuth, K. (2000) Prosodic constraints and the learner's environment: a corpus study. In S. C. Howell, S. A. Fish & T. Keith-Lucas (eds*). Proceedings of the 24th Annual Boston University Conference on Language Development.* Somerville, MA: Cascadilla Press. pp. 597-608

Rosen, S. (2003) Auditory processing in dyslexia and specific language impairment: Is there a deficit? What is its nature? Does it explain anything? *Journal of Phonetics,* 31: 509-527

Rumelhart, D.E. & McClelland, J.L. (1986) On learning the past tenses of English verbs. In J.L. McClelland, D.E. Rumelhart & PDP Research Group (eds) *Parallel distributed processing: explorations in the microstructures of cognition (vol.2).* Cambridge, MA: Bradford/ MIT Press

Sahlen, B., Reuterskiold-Wagner, C., Nettelbladt, U. & Radeborg, K. (1999) Non-word repetition in children with language impairment – pitfalls and possibilities. *International Journal of Language and Communication Disorders,* 34: 337-352

Schwartz, R., Leonard, L., Folger, M.K. & Wilcox, M.J. (1980) Early phonological development in normal-speaking and language disordered children: Evidence for a synergistic view of linguistic disorders. *Journal of Speech and Hearing Disorders*, 45: 357-377

Semel, E., Wiig, E.H. & Secord, W.A. (1995) *Clinical Evaluation of Language Fundamentals 3rd Edition, UK Edition.* London: The Psychological Corporation

SLI Consortium (2002) A Genomewide Scan Identifies Two Novel Loci Involved in Specific Language Impairment. *American Journal of Human Genetics*, 70: 384-398

Smit, A. (1993) Phonologic error distributions in Iowa-Nebraska Articulation Norms Project: word-initial consonant clusters. *Journal of Speech and Hearing Research*, 36: 931-947

Smith, N.V. (1999) *Chomsky: Ideas and Ideals.* Cambridge, UK: Cambridge University Press

Snowling, M. (2000) *Dyslexia.* 2nd Edition. Oxford, UK: Blackwell Publishers

Snowling, M., Chiat, S. & Hulme, C. (1991) Words, non-words and phonological processes: Some comments on Gathercole, Willis, Emslie and Baddeley. *Applied Psycholinguistics*, 12: 369-373

Spencer, A. (1991) *Morphological theory: An introduction to word structure in generative grammar.* Oxford: Blackwell

Stavrakaki, S. (2001) Comprehension of reversible relative clauses in specifically language impaired and normal Greek children. *Brain and Language*, 77: 419-431

Stavrakaki, S. & van der Lely, H.K.J. (submitted) Greek SLI children's interpretation of pronouns: evidence for selective deficits in the computational component of language. *Language*

Stemberger, J.P. (1983) Inflectional malapropisms: Form-based errors in English morphology. *Linguistics*, 21: 573-602

Stemberger, J.P. (1993) Vowel dominance in over-regularisation. *Journal of Child Language*, 20: 503-521.

Stemberger, J.P. (1995) Phonological and lexical constraints on morphological processing. In L. Feldman (ed.) *Morphological aspects of processing.* Amsterdam: North Holland. pp. 247-267

Stemberger, J.P. & Bernhardt, B.H. (1997) Phonological constraints and morphological development. In E. Hughes, M. Hughes & A, Greenhill (eds). *Proceedings of the 24th Annual Boston University Conference on Language Development.* Somerville, MA: Cascadilla Press. pp. 602-614

Stemberger, J.P. & Bernhardt, B.H. (1999) The emergence of faithfulness. In B. MacWhinney (ed.), *The emergence of language.* Mahweh, NJ: Erlbaum. pp. 417-446

Stemberger, J.P. & MacWhinney, B. (1986a) Frequency and the lexical storage of regularly inflected forms. *Memory and Cognition*, 14 :17-26

Stemberger, J.P. & MacWhinney, B. (1986b) Form-oriented inflectional errors in language processing. *Cognitive Psychology*, 18: 329-354

Stevens, T. & Karmiloff-Smith, A. (1997) Word-learning in a special population: do individuals with Williams Syndrome obey lexical constraints? *Journal of Child Language*, 24: 737-765

Taatgen, N.A. & Anderson, J.R. (2002) Why do children say 'broke'? A model for learning the past tense without feedback. *Cognition*, 86: 123-155

Tager-Flusberg, H. (ed.) (1999) *Neurodevelopmental disorders*. Cambridge, Mass: MIT Press

Tallal, P. & Piercy, M. (1973a) Defects of non-verbal auditory perception in children with developmental aphasia. *Nature*, 241: 468-469

Tallal, P. & Piercy, M. (1973b) Developmental aphasia: Impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologia*, 11: 389-398

Thomas, M.S.C., Grant, J., Barham, Z., Gsodl, M., Laing, E., Lakusta, L., Tyler, L.K., Grice, S., Paterson, S. & Karmiloff-smith, A. (2001) Past tense formation in Williams Syndrome. *Language and Cognitive Processes*, 16: 143 - 176

Thornton, R. & Tesan, G. (2003) Overspecification in Early English. Paper presented at the 5th Generative Approaches to Language Acquisition Conference, 4 - 6th September 2003

Tomblin, J. B. & Pandich, J. (1999) Lessons from children with specific language impairment. *Trends in Cognitive Sciences*, 3: 283-285

Tyler, L.K., Randall, B. & Marslen-Wilson, W.D. (2002) Phonology and neuropsychology of the English past tense. *Neuropsychologia*, 40: 1154-1166

Tzakosta, M. (in prep.) *Acquiring stress in an unknown world: The acquisition of stress in Greek and its implications for Universal Grammar.* PhD dissertation, University of Leiden

Ullman, M. (1999) Acceptability ratings of regular and irregular past tense forms: Evidence for a dual-system model of language from word-frequency and phonological neighbourhood effects. *Language and Cognitive Processes*, 14: 47-67

Ullman, M. (2001) A neuro-cognitive perspective on language: The declarative/ procedural model. *Nature Reviews Neuroscience*, 2: 717-726

Ullman, M.T. & Gopnik, M. (1999) Inflectional morphology in a family with inherited SLI. *Applied Psycholinguistics*, 20: 51-117

van der Lely, H.K.J. (1996) Specifically language impaired and normally developing children: verbal passive versus adjectival passive sentence interpretation. *Lingua*, 98: 243-272 van der Lely, H.K.J. (1996). *Test of Active and Passive Sentences (TAPS).* Available from the author at the Centre for Developmental Language Disorders and Cognitive Neuroscience, University College London, London, UK

van der Lely, H.K.J. (1997a) Narrative discourse in Grammatical specific language impaired children: a modular language deficit? *Journal of Child Language*, 24: 221-256

van der Lely, H.K.J. (1997b) Language and cognitive development in a grammatical SLI boy: modularity and innateness. *Journal of Neurolinguistics*, 10: 75-107

van der Lely, H.K.J. (1997c). *Advanced-Syntactic test of Pronominal reference (A-STOP).* Available from the author at the Centre for Developmental Language Disorders and Cognitive Neuroscience, University College London, London, UK

van der Lely, H.K.J. (1998) SLI in children: Movement, economy and deficits in the computational-syntactic system. *Language Acquisition*, 7: 161-192

van der Lely, H.K.J. (1999) Learning from Grammatical SLI: Response to J.B. Tomblin & J. Pandich. *Trends in Cognitive Sciences,* 3: 286-288

van der Lely, H.K.J. (2000) *Verb Agreement and Tense Test (VATT).* Available from the author at the Centre for Developmental Language Disorders and Cognitive Neuroscience, University College London, London, UK

van der Lely, H.K.J. (2003) Do heterogeneous deficits require heterogeneous theories? SLI subgroups and the RDDR hypothesis. In Y. Levy & J. Schaeffer (eds) *Language competence across populations: Towards a definition of specific language impairment.* Mahwah, NJ: Lawrence Erlbaum. pp. 95-108

van der Lely, H.K.J. & Battell, J. (2003) Wh-movement in children with Grammatical SLI: A test of the RDDR hypothesis. *Language*: 79: 153-181

van der Lely, H.K.J. & Christian, V. (2000) Lexical word formation in children with grammatical SLI: a grammar-specific versus an input-processing deficit? *Cognition*, 75: 33-63

van der Lely, H.K.J. & Harris, J. (1999) *Test of Phonological Structure (TOPhS).* Available from the authors at the Centre for Developmental Language Disorders and Cognitive Neuroscience, University College London, London, UK

van der Lely, H.K.J. & Hennessey, S. (1999) Linguistic determinism and theory of mind: Insight from children with SLI. The 24th Boston University Conference on Language Development, Boston, MA. November 1999

van der Lely, H.K.J. & Howard, D. (1993) Children with specific language impairment: Linguistic impairment or short-term memory deficit? *Journal of Speech and Hearing Research*, 36: 1193-1207 van der Lely, H.K.J., Rosen, S. & Adlard, A. (in press) Grammatical language impairment and the specificity of cognitive domains: Relations between auditory and language abilities. *Cognition*

van der Lely, H.K.J., Rosen, S. & McClelland, A. (1998) Evidence for a grammar-specific deficit in children. *Current Biology*, 8: 1253-1258

van der Lely, H.K.J. & Stollwerck, L. (1996) A grammatical specific language in children: An autosomal dominant inheritance? *Brain and Language*, 52: 484-504

van der Lely, H.K.J. & Stollwerck, L. (1997) Binding theory and grammatical specific language impairment in children. *Cognition*, 62: 245-290

van der Lely, H.K.J & Ullman, M. (1996) The computation and representation of past tense morphology in normally developing and specifically language impaired children. In A. Stringfellow, D. Cahan-Amitay, E. Hughes & A Zukowski (eds) *Proceedings of the 20th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press. pp. 816-827

van der Lely, H.K.J. & Ullman, M. (2001) Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes*, 16: 177-217

Vargha-Khadem, F., Watkins, K., Fletcher, P. & Passingham, R. (1995) Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proceedings National Academy of Science*, 92: 930-933

Wauquier-Gravelines, S., Jakubowicz, C., Sauzet, P., Durand, C. & Franc, S. (1997) Phonological knowledge in developmental language disorders: 'Liaison enchainee' and derivations in French. In A. Sorace, C. Heycock & R. Shillcock (eds) *Proceedings of the GALA 1997 Conference on Language Acquisition*. Human Communication Research Centre, University of Edinburgh. pp. 294-299

Weismer, S.E. & Hesketh, L.J. (1996) Lexical learning by children with Specific Language Impairment: Effects of linguistic input presented at varying speaking rates. *Journal of Speech and Hearing Research*, 39: 177-190

Wexler, K., Schutze, C.T. & Rice, M. (1998) Subject case in children with SLI and unaffected controls: Evidence for the Agr/Tns omission model. *Language Acquisition*, 7: 317-344

Yip, M. (1992) Prosodic morphology in four Chinese dialects. *Journal of East Asian Linguistics*, 1: 1-35

APPENDICES

Appendix A. (Chapter 4)

A.1. Verb stimuli

Condition	Past tense form of stimulus	CO-BUILD frequency*	Mean CO-BUILD frequency*	Francis & Kucera frequency**	Mean Francis & Kucera frequency**
VC-D illegal	*hugged*	0.693	1.175	1.099	1.425
	hummed	0.693		1.099	
	robbed	0.693		1.099	
	fished	1.099		0.000	
	buzzed	0.693		1.099	
	touched	2.197		3.219	
	judged	1.386		1.386	
	washed	1.946		2.398	
VC-D legal	*yelled*	1.099	1.647	3.045	2.089
	wrapped	1.792		1.099	
	tossed	1.099		3.091	
	kissed	1.946		2.773	
	killed	2.996		3.526	
	packed	1.609		2.079	
	coughed	0.693		1.099	
	hopped	1.946		1.792	

Frequencies are calculated as ln (raw frequency+1). *Baayen, Piepenbrock & van Rijn (1993). **Francis, W. N. & Kucera, H. (1982)

A.2. Stimulus sentences

The stimulus sentences are presented here in task order.

1) Last week Kipper killed a rat. Every week I kill a rat. Last week I ______.
2) Yesterday Kipper hugged a friend. Everyday I hug a friend. Yesterday I ______.
3) Yesterday Kipper touched a flower. Everyday I touch a flower. Yesterday I ______.
4) Yesterday Kipper kissed a girl. Everyday I kiss a girl. Yesterday I ______.
5) Yesterday Kipper tossed a pancake. Everyday I toss a pancake. Yesterday I ______.
6) Yesterday Kipper packed a lunchbox. Everyday I pack a lunchbox. Yesterday I ______.
7) Yesterday Kipper wrapped a present. Everyday I wrap a present. Yesterday I ______.
8) Last week Kipper robbed a post office. Every week I rob a post office. Last week I ______.
9) Yesterday Kipper buzzed at a bee. Everyday I buzz at a bee. Yesterday I ______.
10) Yesterday Kipper fished in the river. Everyday I fish in the river. Yesterday I ______.
11) Yesterday Kipper washed a blanket. Everyday I wash a blanket. Yesterday I ______.
12) Last year Kipper judged a competition. Every year I judge a competition. Last year I ______.
13) Last winter Kipper coughed a lot. Every winter I cough a lot. Last winter I ______.
14) Yesterday Kipper yelled at his mum. Every day I yell at my mum. Yesterday I ______.
15) Last night Kipper hummed a tune. Every night I hum a tune. Last night I ______.
16) Last night Kipper hopped around the bed. Every night I hop around the bed. Last night I ______.

Appendix B. (Chapter 6)

B.1. Stimulus sentences

Stimulus sentences are presented to the child in pairs. The pairs of sentences are not presented here in task order.

1) Yesterday I *make/ made* a pie.
2) Yesterday I *make/ made* a friend.
3) Yesterday I *make/ made* a jelly.

4) Yesterday I *play/ played* at home.
5) Yesterday I *play/ played* outside.
6) Yesterday I *play/ played* in the rain.

7) Yesterday I *hold/ held* a crown.
8) Yesterday I *hold/ held* a kitten.
9) Yesterday I *hold/ held* a hamster.

10) Yesterday I *yell/ yelled* at mum.
11) Yesterday I *yell/ yelled* at everyone.
12) Yesterday I *yell/ yelled* in anger.

13) Yesterday I *ride/ rode* in a car.
14) Yesterday I *ride/ rode* a horse.
15) Yesterday I *ride/ rode* a donkey.

16) Yesterday I *sew/ sewed* a dress.
17) Yesterday I *sew/ sewed* an apron.
18) Yesterday I *sew/ sewed* a shirt.

19) Yesterday I *hear/ heard* a bang.
20) Yesterday I *hear/ heard* a storm.
21) Yesterday I *hear/ heard* a bird.

22) Yesterday I *purr/ purred* all day.
23) Yesterday I *purr/ purred* a lot.
24) Yesterday I *purr/ purred* all morning.

25) Yesterday I *tell/ told* a joke.
26) Yesterday I *tell/ told* a story.
27) Yesterday I *tell/ told* a lie.

28) Yesterday I *roll/ rolled* out of bed.
29) Yesterday I *roll/ rolled* out the playdoh.
30) Yesterday I *roll/ rolled* in the grass.

31) Yesterday I *sleep/ slept* all day.
32) Yesterday I *sleep/ slept* until late.
33) Yesterday I *sleep/ slept* on the sofa.

34) Yesterday I *step/ stepped* in a puddle.
35) Yesterday I *step/ stepped* in some mud.
36) Yesterday I *step/ stepped* on an ant.

37) Yesterday I *lose/ lost* a bet.
38) Yesterday I *lose/ lost* a sock.
39) Yesterday I *lose/ lost* a book.

40) Yesterday I *toss/ tossed* a pancake.
41) Yesterday I *toss/ tossed* a coin.
42) Yesterday I *toss/ tossed* a ball.

43) Yesterday I *find/ found* a puppy.
44) Yesterday I *find/ found* a penny.
45) Yesterday I *find/ found* a purse.

46) Yesterday I *frown/ frowned* a bit.
47) Yesterday I *frown/ frowned* all day.
48) Yesterday I *frown/ frowned* at everyone

Appendix C. (Chapter 7)

C.1. Verb stimuli

Condition	Past tense form of stimulus	CO-BUILD frequency*	Mean CO-BUILD frequency*	Francis & Kucera frequency**	Mean Francis & Kucera frequency**
VV-D	*poured*	2.079	1.502	3.091	2.419
	weighed	1.099		2.485	
	paid	3.258		3.912	
	sewed	0.000		0.000	
	purred	0.000		0.000	
	showed	3.091		4.934	
	sighed	1.792		3.135	
	lied	0.693		1.792	
VC-D illegal	*hugged*	0.693	1.175	1.099	1.425
	hummed	0.693		1.099	
	robbed	0.693		1.099	
	fished	1.099		0.000	
	buzzed	0.693		1.099	
	touched	2.197		3.219	
	judged	1.386		1.386	
	washed	1.946		2.398	
VC-D legal	*yelled*	1.099	1.647	3.045	2.089
	wrapped	1.792		1.099	
	tossed	1.099		3.091	
	kissed	1.946		2.773	
	killed	2.996		3.526	
	packed	1.609		2.079	
	coughed	0.693		1.099	
	hopped	1.946		1.792	
VVC-D	*pinched*	0.693	0.721	1.099	1.133
	banged	0.693		1.609	
	winked	0.693		2.079	
	solved	1.609		0.693	
	milked	0.000		0.000	
	punched	0.693		0.693	
	munched	0.000		0.693	
	danced	1.386		2.197	

Frequencies are calculated as ln (raw frequency+1). *Baayen, Piepenbrock & van Rijn (1993). **Francis, W. N. & Kucera, H. (1982)

C.2. Stimulus sentences

The stimulus sentences are presented here in task order.

1) Yesterday Kipper winked at his mum. Everyday I wink at my mum. Yesterday I ______.
2) Last week Kipper killed a rat. Every week I kill a rat. Last week I ______.
3) Last year Kipper sewed a dress. Every year I sew a dress. Last year I ______.
4) Last month Kipper solved a crime. Every month I solve a crime. Last month I ______.
5) Yesterday Kipper hugged a friend. Everyday I hug a friend. Yesterday I ______.
6) Yesterday Kipper touched a flower. Everyday I touch a flower. Yesterday I ______.
7) Yesterday Kipper lied a bit. Everyday I lie a bit. Yesterday I ______.

8) Yesterday Kipper kissed a girl. Everyday I kiss a girl. Yesterday I ______.
9) Yesterday Kipper tossed a pancake. Everyday I toss a pancake. Yesterday I ______.
10) Last week Kipper punched a burglar. Every week I punch a burglar. Last week I ______.
11) Last night Kipper poured a drink. Every night I pour a drink. Last night I ______.
12) Yesterday Kipper packed a lunchbox. Everyday I pack a lunchbox. Yesterday I ______.
13) Yesterday Kipper wrapped a present. Everyday I wrap a present. Yesterday I ______.
14) Yesterday Kipper weighed a parcel. Everyday I weigh a parcel. Yesterday I ______.
15) Last night Kipper danced a jig. Every night I dance a jig. Last night I ______.
16) Last week Kipper robbed a post office. Every week I rob a post office. Last week I ______.
17) Yesterday Kipper buzzed at a bee. Everyday I buzz at a bee. Yesterday I ______.
18) Yesterday Kipper fished in the river. Everyday I fish in the river. Yesterday I ______.
19) Last week Kipper paid a bill. Every week I pay a bill. Last week I ______.
20) Yesterday Kipper washed a blanket. Everyday I wash a blanket. Yesterday I ______.
21) Yesterday Kipper pinched a hamster. Everyday I pinch a hamster. Yesterday I ______.
22) Yesterday Kipper showed off to his friends. Everyday I show off to my friends. Yesterday I ______.
23) Yesterday Kipper munched a carrot. Every day I munch a carrot. Yesterday I ______.
24) Last year Kipper judged a competition. Every year I judge a competition. Last year I ______.
25) Yesterday Kipper sighed a bit. Everyday I sigh a bit. Yesterday I ______.
26) Last winter Kipper coughed a lot. Every winter I cough a lot. Last winter I ______.
27) Yesterday Kipper purred at a cat. Everyday I purr at a cat. Yesterday I ______.
28) Last night Kipper banged on the door. Every night I bang on the door. Last night I ______.
29) Yesterday Kipper yelled at his mum. Every day I yell at my mum. Yesterday I ______.
30) Last night Kipper hummed a tune. Every night I hum a tune. Last night I ______.
31) Last night Kipper hopped around the bed. Every night I hop around the bed. Last night I ______.
32) Yesterday Kipper milked a cow. Every day I milk a cow. Yesterday I ______.

Appendix D. (Chapter 8)

D.1. Verb stimuli

Condition		Stimulus	Frequency*	Mean frequency*
d/t-ɪd	Regular	*Sort*	2.398	3.920
		Rent	3.258	
		Rest	4.357	
		Lift	4.248	
		Need	6.026	
		Melt	3.497	
		Start	4.635	
		plant	2.944	
VV-D	Regular	*pay*	5.787	3.815
		tie	3.932	
		sew	2.944	
		pour	3.892	
		weigh	3.526	
		chew	2.833	
		row	1.792	
		play	5.811	
SS-D	Regular	*whisper*	3.466	3.935
		carry	5.720	
		follow	6.293	
		tickle	1.099	
		answer	4.898	
		marry	4.875	
		whistle	2.585	
		empty	2.585	
d/t-ɪd	Irregular	*lead*	4.419	3.769
		bite	2.079	
		shoot	2.944	
		fight	3.178	
		get	5.823	
		read	3.611	
		ride	3.714	
		meet	4.382	
VV-D	Irregular	*blow*	2.585	3.860
		grow	4.190	
		wear	4.190	
		tear	2.773	
		draw	4.159	
		fly	3.332	
		throw	3.850	
		see	5.823	
VC-D	Irregular	*break*	4.205	3.846
		shake	4.060	
		steal	2.398	
		win	3.829	
		run	4.905	
		give	5.656	
		dig	2.079	
		choose	3.638	

* Frequencies are calculated as ln (raw frequency+1). Francis, W. N. & Kucera, H. (1982)

D.2. Stimulus sentences

The stimulus sentences are presented here in task order.

1) Everyday I get a present. Yesterday I ________.
2) Everyday I weigh a parcel. Yesterday I ________.
3) Everyday I tear a newspaper. Yesterday I ________.
4) Everyday I sort the washing. Yesterday I ________.
5) Everyday I pour a drink. Yesterday I ________.
6) Everyday I run a race. Yesterday I ________.
7) Everyday I rent a video. Yesterday I ________.
8) Everyday I whistle a tune. Yesterday I ________.
9) Everyday I lead a donkey. Yesterday I ________.
10) Everyday I dig a hole. Yesterday I ________.
11) Everyday I empty a dustbin. Yesterday I ________.
12) Everyday I plant a tree. Yesterday I ________.
13) Everyday I tickle a dog. Yesterday I ________.
14) Everyday I fly a kite. Yesterday I ________.
15) Everyday I bite an apple. Yesterday I ________.
16) Everyday I play a game. Yesterday I ________.
17) Everyday I draw a picture. Yesterday I ________.
18) Everyday I start a fire. Yesterday I ________.
19) Everyday I choose a present. Yesterday I ________.
20) Everyday I read a magazine. Yesterday I ________.
21) Everyday I pay a man. Yesterday I ________.
22) Everyday I win a prize. Yesterday I ________.
23) Everyday I shoot a pigeon. Yesterday I ________.
24) Everyday I sew a dress. Yesterday I ________.
25) Everyday I follow a zebra. Yesterday I ________.
26) Everyday I break a plate. Yesterday I ________.
27) Everyday I blow a bubble. Yesterday I ________.
28) Everyday I rest a while. Yesterday I ________.
29) Everyday I fight an alien. Yesterday I ________.
30) Everyday I need a drink. Yesterday I ________.
31) Everyday I give a speech. Yesterday I ________.
32) Everyday I wear a crown. Yesterday I ________.
33) Everyday I chew a toffee. Yesterday I ________.
34) Everyday I tie a bow. Yesterday I ________.
35) Everyday I meet a princess. Yesterday I ________.
36) Everyday I throw a ball. Yesterday I ________.
37) Everyday I carry a puppy. Yesterday I ________.
38) Everyday I marry a dancer. Yesterday I ________.
39) Everyday I melt a snowman. Yesterday I ________.
40) Everyday I see a pirate. Yesterday I ________.
41) Everyday I whisper a secret. Yesterday I ________.
42) Everyday I lift a log. Yesterday I ________.
43) Everyday I steal a watch. Yesterday I ________.
44) Everyday I row a boat. Yesterday I ________.
45) Everyday I answer a question. Yesterday I ________.
46) Everyday I shake a leg. Yesterday I ________.
47) Everyday I ride a bike. Yesterday I ________.
48) Everyday I grow a beard. Yesterday I ________.

Appendix E. (Chapter 9)

E.1. Noun stimuli

Condition	Stimulus	Familiarity	Complexity	Age of acquisition	Frequency*
s-ø	*dog*	3.47	3.80	169	4.990
	boot	2.07	3.00	251	3.401
	duck	2.67	3.60	164	1.792
	fish	2.93	3.13		3.497
	cup	2.67	2.47		4.060
	bed	3.73	2.67	169	4.934
	comb	2.93	2.73		1.792
	cake	3.93	2.73	214	2.773
	knife	2.72	2.31		4.454
	goat	2.07	4.00		2.079
	sock	2.40	1.80	172	2.303
	pig	1.80	3.67	233	2.639
Mean		**2.78**	**2.99**		**3.226**
sw-ø	*carrot*	3.07	1.93		1.609
	necklace	2.80	1.73		1.609
	lemon	2.20	2.20	280	2.773
	basket	2.52	2.72		2.944
	orange	2.79	1.90	203	2.708
	parrot	3.00	4.17		0.693
	rocket	2.50	3.29		3.091
	tortoise	2.73	3.33		1.386
	lettuce	2.43	3.38		0.000
	rabbit	3.40	3.93	206	2.773
	mountain	2.03	2.52	283	4.585
	lion	2.00	3.93	244	3.258
Mean		**2.62**	**2.92**		**2.286**
3σ- ø	*butterfly*				
	motorbike				
	crocodile				
	dinosaur				
	elephant	**Note:** This is condition is not used for the phonological analyses, and therefore items are not matched for familiarity or complexity. Nor are the age of acquisition and frequency relevant.			
	envelope				
	ladybird				
	pelican				
	pyramid				
	telephone				
	tricycle				
	saxophone				
s-z	*bears*	2.50	3.93		3.178
	bees	1.93	4.13		3.296
	bows	2.87	2.53	271	2.565
	cows	2.70	3.89		3.829
	doors	2.60	2.80	214	5.852
	keys	3.33	3.33		4.263
	pears	2.80	1.93		2.079
	bells	2.67	2.67		3.135

	balls	3.53	3.53	150	4.812
	ties	2.33	2.33		3.296
	chairs	2.87	2.87		4.489
	jars	2.10	2.10	242	2.944
Mean		**2.69**	**2.96**		**3.649**
sw-*z*	*anchors*	2.14	3.10		2.833
	whistles	2.72	3.21		1.099
	apples	3.20	2.33	211	2.708
	candles	3.27	2.20		3.135
	hammers	2.64	3.07	278	1.792
	ladders	2.27	2.13		2.944
	tables	2.53	1.67		5.489
	tigers	2.21	4.57		2.197
	zebras	2.38	4.17		0.000
	feathers	2.47	2.53		2.944
	bottles	2.53	2.33		4.500
	buttons	2.93	1.20	192	2.996
Mean		**2.61**	**2.71**		**2.720**
s-*ɪz*	*axes*	1.96	2.19	311	2.944
	buses	3.33	3.47		3.738
	boxes	2.20	1.60	192	4.407
	horses	3.53	3.53		5.313
	vases	2.50	3.21	297	2.708
	foxes	2.21	3.79	283	2.398
	watches	3.53	2.93		3.434
	benches	2.53	2.80	311	3.738
	churches	2.38	3.41	278	6.111
	torches	3.57	3.36		1.386
	fences	1.55	1.86		3.829
	purses	n/a	n/a		2.708
Mean		**2.66**	**2.92**		**3.560**

* Frequencies are calculated as ln (raw frequency+1). Francis, W. N. & Kucera, H. (1982)

E.2. Verb stimuli

Condition	**Inflected form of stimulus**	**Frequency***	**Mean frequency***
s-*ing*	hiding	3.555	2.805
	dancing	3.850	
	bouncing	3.611	
	wrapping	2.197	
	diving	1.946	
	mending	1.792	
	chasing	1.946	
	leading	5.011	
	chopping	2.197	
	sipping	1.945	
sw-*ing*	whispering	2.197	2.904
	watering	1.792	
	measuring	4.025	
	emptying	2.639	
	hammering	0.693	

	carrying	5.063	
	following	5.670	
	finishing	3.466	
	galloping	1.099	
	balancing	2.398	
ws-ing	embracing	2.485	2.935
	directing	3.091	
	applauding	2.079	
	arranging	2.833	
	adjusting	2.833	
	relaxing	3.219	
	deciding	3.970	
	repairing	1.946	
	returning	4.700	
	arresting	2.197	

* Frequencies are calculated as ln (raw frequency+1). Francis, W. N. & Kucera, H. (1982)

E.3. Stimulus sentences

The stimulus sentences are presented here in task order.

1) This girl likes to hide behind the tree. Tell me what she is doing.
2) This man likes to relax on the sofa. Tell me what he is doing.
3) This girl likes to dance to the music. Tell me what she is doing.
4) This lady needs to water her plant. Tell me what she is doing.
5) This policeman has to arrest the man. Tell me what he is doing.
6) This boy likes to bounce the ball. Tell me what he is doing.
7) This girl has to wrap a present. Tell me what she is doing.
8) This woman needs to adjust her hat. Tell me what she is doing.
9) This woman needs to arrange the flowers. Tell me what she is doing.
10) This lady likes to whisper to her friend. Tell me what she is doing.
11) This lady likes to dive in the pool. Tell me what she is doing.
12) This girl likes to applaud her friends. Tell me what she is doing.
13) This doctor has to measure the boy. Tell me what he is doing.
14) This girl has to empty the rubbish bin. Tell me what she is doing.
15) This man needs to hammer the nail. Tell me what he is doing.
16) This man needs to decide what to do. Tell me what he is doing.
17) This man has to carry the sack. Tell me what he is doing.
18) This man likes to return to his house. Tell me what he is doing.
19) This man's job is to mend the shoe. Tell me what he is doing.
20) This dog likes to chase the boy. Tell me what it is doing.
21) This man has to repair the car. Tell me what he is doing.
22) This girl needs to lead the horse. Tell me what she is doing.
23) This boy has to follow the footprints. Tell me what he is doing.
24) This girl likes to embrace her mother. Tell me what she is doing.
25) This policeman has to direct the traffic. Tell me what he is doing.
26) These runners want to finish the race. Tell me what they are doing.
27) This man has to chop the wood. Tell me what he is doing.
28) This girl likes to gallop on her horse. Tell me what she is doing.
29) This boy likes to balance on his head. Tell me what he is doing.
30) This boy likes to sip his drink. Tell me what he is doing.

Appendix F. (Chapter 10)

F.1. Adjectival stimuli

Condition	Raw frequency*			Condition	Raw frequency*		
1 syllable	stem	-er	-est	2 syllables	stem	-er	-est
big	359	34	24	*narrow*	63	7	0
fat	47	3	0	*muddy*	10	0	0
hard	140	14	9	*hairy*	5	0	0
tall	55	7	0	*happy*	97	11	3
sad	35	1	0	*silly*	15	0	1
short	195	18	3	*curly*	5	0	0
dark	149	2	2	*tidy*	1	0	0
wide	118	17	3	*dirty*	36	0	0
red	169	3	0	*heavy*	110	14	2
thin	90	6	0	*funny*	40	1	2
Mean	**135.7**	**10.5**	**4.1**	**Mean**	**38.2**	**3.3**	**0.8**

* Francis, W. N. & Kucera, H. (1982)

Appendix G. (Chapter 11)

G.1. Noun stimuli

Condition		Condition	
Singular	Raw frequency* of derived adjective	Plural	Raw frequency* of derived adjective
sun	12	*spots*	1
mud	10	*holes*	0
wool	3	*waves*	2
sand	6	*curls*	5
soap	1	*frills*	1
dirt	36	*scales*	0
hair	5	*stars*	0
rain	4	*clouds*	2
juice	6	*rocks*	9
fur	0	*stripes*	0
Mean	**8.3**	**Mean**	**2.0**

* Francis, W. N. & Kucera, H. (1982)